EL FENÓMENO DE LA VIDA



Francisco Varela



EL FENÓMENO DE LA VIDA

FRANCISCO VARELA

DOLMEN EDICIONES
Caracas · Montevideo · Santiago de Chile

Para la presente edición:
©Dolmen Ediciones S. A.
Cirujano Guzmán 194, Providencia, Santiago de Chile

Nº Inscripción: 116.301
ISBN: 956-201-456-8

Dirección editorial: Cristóbal Santa Cruz
Diseño de portada y diagramación: Job López
Traducción: Traduki
Revisión de traducción: Jaime Cordero, Cristóbal Santa Cruz

Esta edición de 1000 ejemplares
se terminó de imprimir en octubre del 2000,
en Dolmen Ediciones S.A., Santiago de Chile.



Impreso en Chile/Printed in Chile

INDICE

2. *Organismo*

II
MENTE

1. *Cognición*

III
EPISTEMOLOGÍA

IV
Discusiones

INTRODUCCIÓN DEL AUTOR

"Sólo la vida puede conocer la vida."
Hans Jonas

Escoger una selección de ensayos representativos de treinta años de trabajo científico y reflexión sobre la vida y el conocer es una empresa literalmente imposible. Al contemplar el conjunto aquí seleccionado, listo para entrar a imprenta, me invade el sentimiento de no haber podido incluir tantas otras cosas que parecen de pronto esenciales, una frustración de saber esta selección incompleta e impresionista. Al mismo tiempo me sobrecoge la impresión sobria y descarnada de lo inadecuado de la expresión, de lo efímero y circunstancial de los textos, en breve, de la pobreza de lo dicho frente a lo que se quiere decir o lo que se pudo haber dicho.

Con todo y a pesar de todo, estos textos reflejan un itinerario forjado, una perspectiva trazada a punta de volver una y otra vez a ciertas cuestiones fundamentales, como la autonomía y la individualidad de lo vivo, el origen del significado y el conocimiento como emergencia encarnada, la conciencia en tanto experiencia vivida y realización biológica en un todo indisoluble. El corazón de estas ideas está contenido en las secciones 'Organismo' y 'Mente', a las que las otras secciones, 'Epistemología' y 'Discusiones' sirven de apoyo.

Quiero disipar desde el comienzo la idea de que se trata de un sistema o teoría unificada sobre la vida y sus avatares. Tengo una desconfianza instintiva respecto de todo sistema globalizador que busque explicar de manera más o menos cerrada lo que aquí es más bien terreno de indagación, de exploración siempre renovada. En consecuencia, muchos de los ensayos se repiten en

cierto grado, las ideas se reencuentran en reflejos cruzados sobre otros aspectos del mismo centro. Que el lector tenga paciencia de recorer este camino de *corsi e ricorsi*, que atrapa a la pasada un estilo de trabajo, no un sistema explicativo sino una instantánea de un arco de movimiento en plena expresión.

Lo que me interesa es el fenómeno de la vida en toda su gloria y majestad. Esto quiere decir evitar de caer en la tentación dominante de reducir el fenómeno de la vida a la sorpresa de ver que surge de un mundo material molecular muerto y sin significación. ¿Cómo puede ser posible tal fenómeno? Curiosa sorpresa de la modernidad que parte del mundo impersonal y muerto del universo físico, y que se ve forzada, casi a contrapelo, a rescatar lo que lo contradice centralmente, puesto que su especificidad es lo vivo significante. Gran parte de las ideas que presento aquí son precisamente el esfuerzo de romper ese sueño newtoniano y de recentrar el fenómeno de la vida sobre sus propias bases, es decir, no como evento improbable que escapa a la muerte física, sino como ocurrencia central de la naturaleza que funda todo conocer, y en particular el conocimiento del mundo físico. Es precisamente porque este punto de vista va en gran medida a contrapelo del pesamiento científico dominante que estas ideas han avanzado a tientas, sin muchos antecedentes, y en general en un ambiente de desconfianza o de franco rechazo. Se trata de un pensamiento que molesta porque pone en evidencia las limitaciones del fisicalismo, base conceptual de la filosofía espontánea de la biología hoy día respecto de la comprensión de lo vivo y de la conciencia.

¿Y qué reemplaza este fisicalismo dominante? La conclusión a la que estos ensayos apuntan es que lo central es una circularidad inalienable entre el acto de conocer y vivir, entre el universo de lo vivo y el conocer como objeto de estudio. Dicho de otra manera, el fenómeno de la vida como un todo quiere decir, precisamente que el acto de vivir precede a la explicación de los orígenes de la vida sobre la tierra. Que el conocer precede a la comprensión del conocer visto como mecanismo biológico y neuronal. Que la experiencia vivida es la base misma de la exploración científica de la conciencia.

Este es, pienso, el sentido del epígrafe de Jonas que abre esta introducción. No sólo es la frase concisa y evocadora, sino que la pongo ahí porque el título de este libro es un intertexto a la obra magistral de Jonas, *The Phenomenon of Life*, que apareciera publicado en 1966. Vine a descubrir a Jonas hace pocos años cuando las ideas de este libro ya estaban asentadas. Con pasión encontré una filosofía de la biología que daba una amplitud a precisamente *todo* el fenómeno de la vida hasta ahora inigualado. Por lo mismo este pensador, todavía poco conocido, me dio una confianza en mis propias intuiciones. Lo traigo aquí como figura tutelar que sigue siendo fuente de inspiración.

Donde Jonas ha sido un precursor, y donde este libro se sitúa en continuidad, se abre hoy en día un enorme espacio de trabajo científico filosófico que toca ya los fundamentos mismos de la ciencia. Así los hilos que se dejan aquí sin entretejer, y las interrogaciones menos elaboradas, son anuncios de lo que podría advenir.

Los ensayos incluidos aquí pertenecen a circunstancias muy variadas: artículos encargados, participación a simposios, introducción a libros colectivos, lecturas de divulgación. Los niveles y tonos son, en consecuencia muy variables y con grados de dificultad desiguales. Algunos de los textos fueron publicados en conjunto con otros colegas (L. Luisi, A. Lazcano, E. Thompson, J. P. Dupuy y J. Shear). En todos ellos fui yo el principal responsable de la redacción, y consintieron generosamente a que fueran incorporados en este libro.

Me queda agradecer a Cristóbal Santa Cruz, cuyo entusiasmo ha sido el gestor de este libro y que además ha traducido buena parte de los textos.

Me queda expresar mi deuda con Amy Cohen-Varela, sin la cual no estaría escribiendo estas palabras.

París, julio 2000

I

ORGANISMO

I

ORGANISMO

1. Lo vivo

¿QUÉ ES LA VIDA?
Definiendo la vida y la transición a la vida

Este texto introduce de manera informal las ideas y los problemas de base que llevan a la cuestión de la definición de la vida y sus orígenes sobre la Tierra. Volveremos sobre los mismos temas en otros artículos de esta Sección ("Patrones de vida" y "Organismo") de maneras diversas, como círculos sobre un núcleo fundamental. Pero quedan aquí esbozados una primera vez en este libro.

1. Introducción

Hace más de cincuenta años, el distinguido físico austriaco Erwin Schrödinger escribió un libro con el mismo título que este trabajo. Sin embargo, incluso antes de esa fecha, el intento por definir la vida se ha topado con actitudes de escepticismo o de condescendencia. Detrás de esta actitud negativa hay un argumento típico: semejante empresa no es ni útil ni fácil, ya que todos sabemos lo que es la vida, pero es imposible traducirlo en palabras. También se ha argumentado que la transición desde la no vida a la vida es un proceso continuo, y que por lo tanto, es imposible discriminar entre lo vivo y lo no vivo.

Este artículo hace frente a estas actitudes negativas e intenta proporcionar una definición científica de lo vivo. Hay unas cuantas

---

(Luisi,L. Lazcano,A. and Varela,F. (1996), Autopoiesis: the very idea, in: *Defining Life: The central problem in theoretical biology*, M.Rizzotti (Ed.), Universitá de Padova, Padova, pp.146-167.)

buenas razones para ello. En primer lugar, nos enfrentamos al desafío de captar en una formulación explícita la cualidad de la vida: ¿cómo expresar aquello que comparten los microorganismos, las plantas, los animales, las callampas, y los hombres y que los diferencia del mundo inanimado de las rocas y las galaxias?

Además, la importancia de dar una definición de lo vivo debiese resultarles obvia a todos aquellos científicos que trabajan en el campo del origen de la vida y en modelos experimentales de células vivas mínimas. A decir verdad, parece difícil llevar a cabo este tipo de investigación sin tener una noción clara acerca de lo que es la vida en su forma mínima. Por extraño que parezca, las ciencias de la vida generalmente no establecen esta definición ni criterio, a pesar de la extensa literatura sobre los orígenes de la vida y los modelos explicativos de la vida.

Una definición de la vida también puede tener impacto en la investigación espacial: si enviamos sondas espaciales a un planeta lejano para buscar vida más allá de los confines de la tierra, ¿qué es lo que deben buscar? Claramente, también en este caso hay que tener algún patrón característico que nos permita reconocer la vida.

¿Cuáles deberían ser las propiedades de una "buena" definición de la vida? Nuestra posición es que ésta debe restringirse a la vida mínima, a la pregunta: ¿cuál es la unidad estructural más sencilla posible que encarna una definición de lo vivo? Al restringirnos a la vida mínima, obviamos el problema que plantean los desarrollos evolutivos posteriores del organismo y la vida multicelular. Esto también implica que en este estadio de la definición dejamos de lado propiedades emergentes tales como la inteligencia, la mente, la conciencia, y otras similares —ciertamente los atributos más maravillosos de la vida humana, pero desde luego no aquellos que simplifican la definición de la vida misma.

Estas puntualizaciones son insuficientes. Necesitamos otra, que en nuestra opinión —como autores y científicos que se interesan por la materia— constituye una de las principales fuentes de confusión e incompresión entre los científicos cuando se trata de definir la vida. Se trata de lo siguiente: hay dos tipos de científicos

que tratan de responder a la pregunta "¿qué es la vida?" con dos mentalidades muy diferentes. Uno de estos grupos está constituido por los biólogos, los genetistas, la mayoría de los biólogos moleculares, así como por los evolucionistas, para quienes la vida constituye principalmente un *fenómeno de población genética*: una generación de *E. coli* da origen a la siguiente generación de *E. coli*; una cultura de arvejas que produce la siguiente cultura; y del mismo modo para todos los animales y plantas. Esta es la vida en términos de continuidad histórica, una continuidad que sólo se ve perturbada a lo largo de una gran extensión de tiempo por diferentes cambios evolutivos que dan origen a novedades biológicas.

Pero hay otra perspectiva de la vida, una con la que están familiarizados los químicos y los físicos y en gran medida las personas que trabajan con la inteligencia artificial, aunque las fronteras no son tan claras. Esta es la perspectiva de la vida a nivel del *individuo, aquí y ahora*. El científico examina un sólo espécimen (por ejemplo, un nuevo robot; o un complejo supramolecular sintético; o un solo espécimen de un nuevo tipo de medusa; o un espécimen de vida presunta en un planeta distante) y formula la pregunta: ¿está viva o no esta entidad? En este caso, el análisis se centra principalmente en un solo organismo sometido a examen y puede que los antecedentes históricos carezcan de importancia (a veces porque son desconocidos o porque no se pueden establecer). Este tipo de perspectiva local, centrada en "el aquí y el ahora", es la que exige un criterio para discriminar entre lo vivo y lo no vivo sobre una base más bien inmediata, sin esperar la reproducción (puede que ese espécimen particular sea estéril) o la evolución.

Todas estas puntualizaciones definen la esencia y los límites de este artículo. En primer lugar, nos restringiremos a la vida mínima; en segundo lugar, consideraremos, sobre todo, la definición de la vida a nivel del individuo, ya que la vida, en términos de la genética de una población, es más bien controversial. Sin embargo, nuestra intención es mostrar que ambos niveles están vinculados. Además, uno de los puntos fuertes de este artículo es que intenta proporcionar una definición de la vida, utilizando argumentos de "sentido común", partiendo del nivel de un lego inteligente que carece de instrumentos

científicos y de un conocimientos de la biología molecular o celular. Luego veremos que esta definición macroscópica y de "sentido común" de lo vivo a nivel de un solo individuo no difiere conceptualmente de la definición que podríamos dar sobre la base de un conocimiento de la biología celular.

2. El nivel macroscópico: Conversación con un extraterrestre

Para acercarnos al problema de la definición de la vida a nivel del individuo proporcionada por un lego inteligente sin conocimientos de biología celular o molecular, valgámonos de la metáfora que ya fuera utilizada por Oparin en otro contexto: supongamos que una creatura inteligente de Marte —un extraterrestre— visita la tierra para investigar las formas de vida que existen en nuestro planeta. Trae consigo una larga lista de cosas terrestres y quiere saber cuáles de ellas están vivas y cuáles no. Se encuentra con un granjero inteligente, pero científicamente ingenuo que, muy rápidamente, separa los ítemes del extraterrestre en dos listas: (Tabla 1) cosas vivas (columna izquierda) y no vivas (columna derecha).

hombre	radio
árbol	automóvil
gusano	computador
mula	robot
mosquito	luna
coral	marea

Tabla 1: Cosas vivas y no vivas

El extraterrestre se muestra sorprendido por la seguridad con que el granjero discrimina entre ambas categorías, y quiere saber cuál es la cualidad que caracteriza a todas las cosas vivas (ítemes de la izquierda), y que está ausente en los de la derecha. También exige que esta cualidad pueda reconocerse en un tiempo corto, ya

que sólo dispone de unos cuantos días libres. Cuando, apuntando a la gallina y a la mula, el granjero dice "movimiento" y "crecimiento", el extraterrestre disiente con la cabeza, ya que ni el árbol ni el coral se mueven, ni dan señales de crecimiento que puedan apreciarse en un tiempo de observación razonablemente largo; a la inversa, un trozo de papel se mueve con el viento y la luna y las mareas crecen en forma periódica. Cuando el granjero se vale del criterio "reacción ante los estímulos", el extraterrestre tampoco se muestra convencido: efectivamente, tanto la gallina como la mula y el hombre reaccionan ante el pinchazo de una aguja, pero la callampa y el árbol permanecen insensibles: y por otro lado, el computador y la radio pierden fácilmente su eficacia ante la interferencia de una aguja.

"Los seres vivos" —agrega el granjero que comienza a mostrar señales de irritación— "son capaces de realizar sus funciones gracias al consumo de alimentos y a la consiguiente producción de energía. La energía es transformada en acción por el propio organismo". Pero el extraterrestre señala que el auto y el robot son capaces de moverse haciendo precisamente lo mismo, transformando la energía en acción. "¡Reproducción!" —exclama el granjero. "¡Todos los ítemes en su lista son capaces de reproducirse! La gallina y el hombre son capaces de reproducirse, pero no así la mula" — replica con sorna el extraterrestre. "Además, los niños humanos y los humanos muy ancianos no se reproducen: ¿acaso no están vivos?" "Bien, pero todos los ítemes de mi lista derivan de estructuras parentales" —dice el granjero. ¡Todas las cosas vivas tienen un ancestro: se originan de un ser vivo del mismo tipo!"

El extraterrestre se queda pensativo, pero no está convencido. "Supón —dice él— que encuentro una cosa aislada que se parece a un coral. No hay forma de saber si tiene ancestro; y no hay forma de saber si tendrá progenie. ¿Cómo puedo determinar que están vivos o no en un tiempo de observación relativamente corto?" Desea un criterio que le permita definir la vida aquí y ahora, sin tener que investigar el pasado o las generaciones futuras. ¿Que tienen en común todos los ítemes de la columna izquierda, a diferencia de los de la columna de la derecha?

El granjero, que ahora ha adoptado una actitud más reflexiva, le indica un árbol y señala que éste pierde sus hojas en invierno, pero que vuelve a generarlas a la primavera siguiente. El árbol se regenera desde su propio interior. Lo mismo ocurre con el pelo de muchos animales: se cae y vuelve a crecer nuevamente ¡y crece gracias a una actividad que ocurre al interior del cuerpo! A partir de su propia experiencia, el granjero también sabe que cuando ayuna, o cuando está enfermo, sus miembros y sus órganos se vuelven un poco más pequeños: sin embargo, apenas comienza a comer nuevamente, sus miembros y sus órganos vuelven a crecer. Nuevamente, este crecimiento es fruto de una actividad que ocurre al interior de su propio cuerpo. Esto también debe ocurrir cuando el período de ayuno es muy corto —en otras palabras, el proceso de destrucción y regeneración de los tejidos debe ser un proceso continuo. Concluye que en todos los animales y plantas, a los que define como vivos, los tejidos sufren un proceso continuo de destrucción y regeneración fruto de algún tipo de actividad en su propio interior. ¡Los organismos vivos son capaces de albergar una actividad que regenera sus propios componentes!

Ahora el granjero se siente satisfecho: ¡finalmente ha logrado articular la cualidad que distingue a lo vivo de lo no vivo! El robot, el computador, la radio, la luna etc. son incapaces de regenerarse desde adentro. Si una parte de la radio se descompone, la propia radio es incapaz de construirla nuevamente. Sin embargo, todos los ítemes del lado izquierdo de la tabla poseen esta cualidad: utilizando la energía externa para mantener su propia estructura, y la capacidad para regenerarse desde el interior de la propia estructura. Parece que ésta es la propiedad de la vida que han estado buscando. El extraterrestre sonríe y traza una figura en el suelo con su lapicero láser (Figura 1). En esta figura, S representa un componente del sistema vivo, que está siendo transformado en un producto P; pero el sistema es capaz de regenerar a S, transformando los alimentos que ingresan al organismo nuevamente en S. Y, de este modo, el granjero y el extraterrestre llegan a la siguiente definición "macroscópica" de la vida: podemos decir que un sistema físico está vivo cuando es capaz de transformar la

materia/energía externa en un proceso interno de automantención y autogeneración.

Esta definición y su correspondiente ilustración aparecen en la Figura 1: La vida sería respetada si el sistema mantiene su identidad, lo que significa que las dos tasas (transformación de A en S y de S en P) son comparables. La Figura 1 también ilustra la cualidad de la autorreproducción: si la tasa de conversión de A en S es mayor que la tasa de transformación de S en P, entonces el organismo crecerá, y finalmente se dividirá en dos. Pero la autorreproducción no es un prerrequisito necesario para la vida. El extraterrestre se siente bastante contento con todo esto: se trata de una definición de la vida *hic et nunc* ("aquí y ahora"), una definición que no requiere la búsqueda o la espera de progenie. Para verificar esta definición sólo se requiere un tiempo de observación finito: el tiempo necesario para la autogeneración.

3. El nivel celular: ingresa la autopoiesis

En la sección anterior, llegamos a una definición de la vida a nivel macroscópico y de sentido común, válida para un solo espécimen. Ahora intentaremos definir la vida en términos del patrón general de los organismos unicelulares, nuevamente a nivel de un solo individuo (la premisa de este artículo). Operacionalmente, necesitamos un microscopio y algunos métodos básicos de análisis bioquímico.

A decir verdad, si nuestros dos personajes, el granjero y el extraterrestre, hubiesen tenido acceso a un microscopio y a la tradición de la biología celular, habrían arribado de inmediato a una solución: todos los seres vivos están constituidos por células, lo que claramanente no es el caso de los ítemes de la derecha en la Tabla 1. Obsérvese que a este nivel no es necesario disponer de información sobre la detallada estructura molecular de los componentes celulares. Por el momento, podemos dejar de lado gran parte de la biología molecular contemporánea.

Debemos estar conscientes de las limitaciones de este criterio científico. En primer lugar, se trata de un criterio experimental;

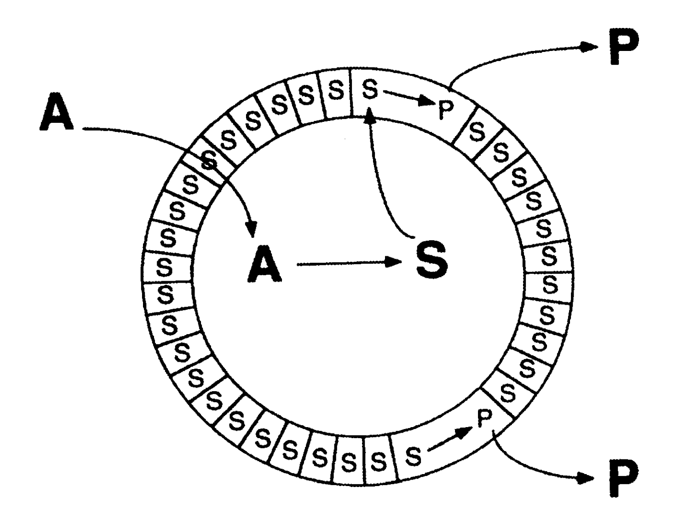


Figura 1: La definición del extraterrestre: este diagrama nos presenta en forma esquemática un organismo constituido por un solo "tejido" S. S se descompone y se transforma en P, pero gracias a la actividad interna del sistema, el nutriente A se transforma nuevamente en el tejido S. Esta representación pictórica corresponde a la siguiente definición de la vida: podemos decir que un sistema físico está vivo si es capaz de transformar la materia/energía externa en un proceso interno de automantención y autogeneración.

es decir, su implementación requiere de una comunidad científica en que estas observaciones adquieren sentido. Va más allá del sentido común ya que supone una comprensión consensual de la vida celular. Además, el criterio se basa en la observación empírica que nunca hemos encontrado vida en la tierra desprovista de células —lo que es un supuesto (ya que pudo haber sido de otra manera durante el origen de la vida misma). Esta definición de la vida también es bastante tautológica: la vida es celular porque no hay vida sin células. Es evidente que para escapar de la tautología tenemos que definir las propiedades básicas de una célula en forma independiente —es decir, sin valernos de las propiedades de la vida. En otras palabras, tenemos que determinar cuáles son las propiedades químicas y físicas necesarias y suficientes para que una célula esté viva: este conjunto

de requisitos mínimos debería, entonces, ser el núcleo de nuestra definición de la vida mínima.

Este salto conceptual fue introducido en los años 70 por H. Maturanay yo mismo con la proposición del concepto de *autopoiesis*, un neologismo acuñado de las raíces griegas para la autoproducción. ¿En que se fundamenta? A partir del esquema de la figura 2, es evidente que una célula se caracteriza, en primer lugar, por una barrera o membrana semipermeable que establece un límite de difusión y permeabilidad que discrimina entre un interior químico (o *sí mismo*) y los medios libremente difusivos en el ambiente externo (*no-sí mismo*). Dentro de este límite, la vida celular es una red metabólica: basada parcialmente en nutrientes que ingresan desde el medio exterior, una celula se mantiene mediante una red de transformaciones químicas. Pero —y este es el punto

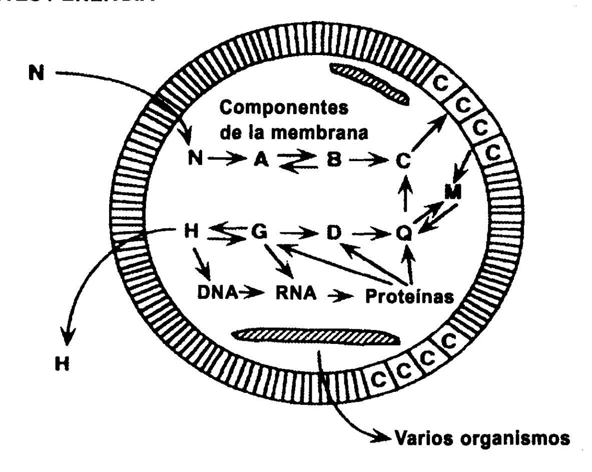


Figura 2: Ilustración esquemática de la lógica celular básica. Esta sencilla figura permite captar el patrón básico de la vida mínima, que se define del siguiente modo: (i) la vida es producto de una organización sistémica emergente, más que de una determinada estructura o reacción molecular; (ii) la vida está ligada a la noción de automantención y de autorregeneración de los componentes del sistema; (iii) la barrera es de vital importancia para discriminar entre el *sí mismo* y el *no-sí mismo*, dando origen a un microambiente en el que puede ocurrir una determinada red de reacciones.

clave— la red de reacciones es capaz de regenerar los componentes que están siendo transformados, incluyendo aquellos que constituyen la barrera o membrana.

En otras palabras, la célula es capaz de automantención gracias a un patrón constitutivo o proceso de generación circular que reemplaza continuamente a los componentes que están siendo destruidos y recrea las condiciones para discriminar entre el sí mismo y el no-sí mismo. Claramente, esta pauta de organización básica podría presentarse en diferentes tipos de unidades, no exclusivamente en una célula básica tal como las conocemos. Esta organización mínima es lo que con Maturana denominamos una unidad autopoiética, la que se define más precisamente del siguiente modo: *un* sistema autopoiético *está organizado (esto es, se define como una unidad) como una red de procesos de producción (síntesis y destrucción) de componentes., en forma tal que estos componentes: (i) se regeneran continuamente e integran la red de transformaciones que los produjo, y (ii) constituyen al sistema como una unidad distinguible en su dominio de existencia.*

Por supuesto, en una célula biológica estas reacciones no son procesos espontáneos, que ocurren al azar —en realidad, son controlados por ácidos nucleicos y las correspondientes enzimas nucleicas ácidas. Sin embargo, este importante punto no modifica la esencia del patrón general que acabamos de describir, en el sentido que la descripción autopoiética de la vida mínima no depende de la estructura. Es de índole general, también podría aplicarse a una "vida no codificada"; por ejemplo a un sistema de vida artificial o a un sistema químico sintético desprovisto de ácidos nucleicos. En otras palabras, si bien la descripción autopoiética de la vida mínima se basa en una célula viva, define una pauta general de la cual la vida basada en un código ADN es tan sólo una expresión —aunque ciertamente la más relevante— para nuestra vida en la Tierra.

Nuevamente, podemos esquematizar todo lo anterior bajo la forma de una figura, y presentar el modelo más sencillo de una unidad autopiética, como el que aparece en la Figura 3. Se trata

de un sistema constituido por una sola barrera constituida por un solo componente S, y conteniendo sólo la reacción que lleva a la síntesis de S y a su descomposición. Dependiendo de la tasa relativa de formación y descomposición de S, la unidad autopoiética puede sufrir varios cambios kinéticos como, por ejemplo, la *autorreproducción*. Pero la pauta más general y más genuina es la de la automantención y la autogeneración. Obsérvese que formalmente la Figura 3 es similar a la sencilla Figura 1 del extraterrestre, pero esta vez valiéndose del conocimiento de la autoformación de una unidad provista de una barrera a la que hemos llegado en esta etapa del análisis.

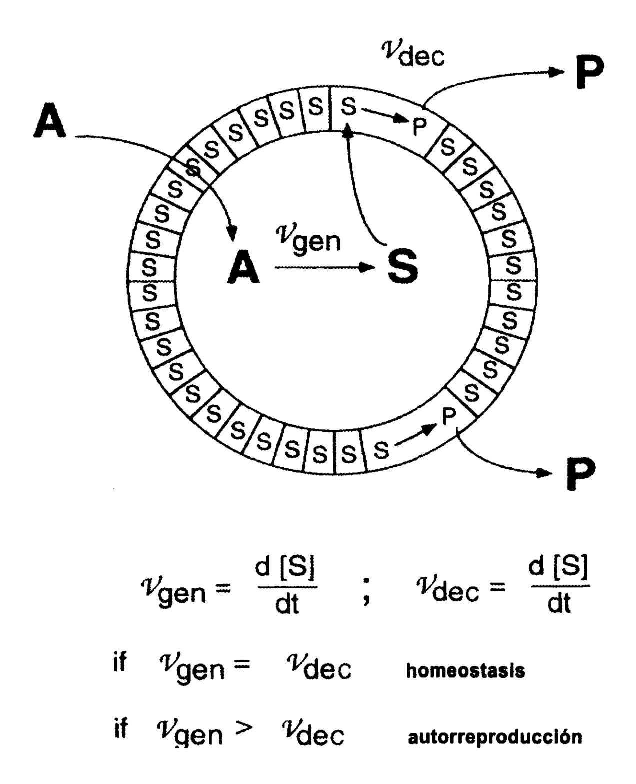


Figura 3: Esquema de la unidad autopoiética más sencilla posible. En este caso, en contradicción con la Figura 1 del extraterrestre, la barrera es una estructura *físicamente especificada*, es decir, es una membrana semipermeable o barrera. La figura 3 puede considerarse como una sobresimplificación de la conducta celular ilustrada en la Figura 2.

Volvamos a repetirlo: la definición de autopoiesis define el esquema general de la vida sin hacer referencia alguna a la estructura de los componentes. El mundo del ADN, del ARN y de las proteínas, que es tan básico para la vida tal como la conocemos en la Tierra, es tan sólo una implementación particular de esta organización autopoiética general.

La organización operacional de una unidad autopoiética es cíclica: los componentes moleculares que produce la red de reacciones internas se reúnen para constituir la unidad provista de una barrera semipermeable que determina un microambiente y la red de reacciones que vuelve a producir los componentes moleculares que... y así sucesivamente.

Esta unidad operacional autopoiética general aparece ilustrada en la Figura 4a, sin perjuicio del tipo de reacciones y de la naturaleza química de los componentes. En contraste, en la figura 4b la función que desempeñan las reacciones desencadenadas por el ADN se proporciona en forma explícita.

Sorprendentemente, hemos llegado a una definición que, al menos superficialmente, se aproxima a las conclusiones macroscópicas del extraterrestre. Sin embargo, en esta definición científica hay elementos nuevos y claves que están lejos de ser de sentido común. De vital importancia es el papel que desempeña la autodefinición de una unidad, que hace de la vida una propiedad emergente y no simplemente un agregado de cualidades. Lo que quiero decir con "emergente" es que la vida mínima puede ser explícitamente descrita como el resultado de una organización, no como el resultado de una reacción de determinados componentes. Esta organización define a la unidad: se trata realmente de una propiedad distribuida de la interacción entre la red metabólica y la membrana, donde las propiedades globales como la difusión determinan las reacciones químicas locales; sin embargo, son las reacciones químicas las que posibilitan la existencia del límite o barrera. Nada de esto es de sentido común y ni siquiera corresponde al pensamiento científico clásico, luego no podemos culpar al granjero por no dar de inmediato con esta definición.

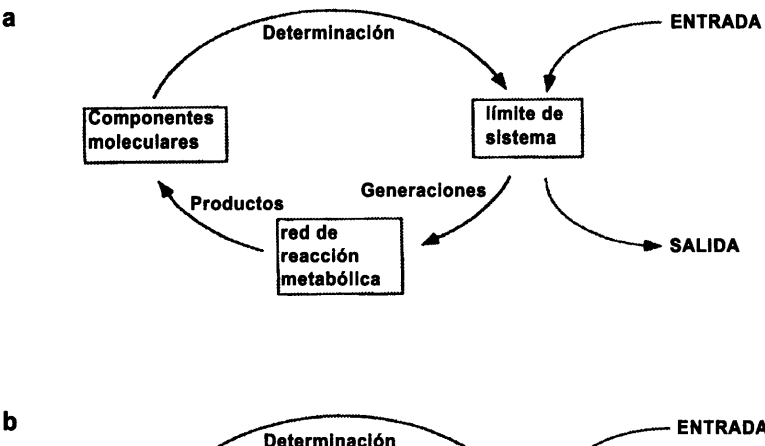
¿Cómo se valida la autopoiesis? ¿Cuándo podemos decir que un sistema es autopoiético, y que constituye, por lo tanto, un ejemplo de vida mínima? La ventaja de una definición explícita es que ahora podemos proporcionar ciertos criterios de validación, que aquí resumiremos en tres:

1. Borde semipermeable: Verifique que el sistema se define por un límite o borde semipermeable constituido por componentes moleculares que permite discriminar entre el interior y el exterior del sistema en relación con los componentes relevantes del sistema. Si es así, proceda con el criterio 2.
2. Red de reacciones: Verifique que los componentes de la barrera son producto de una red de reacciones que opera al interior de la barrera. Si la respuesta es afirmativa, proceda con el criterio 3.
3. Interdependencia: Verifique que la red de reacciones es regenerada por condiciones producidas por la existencia de la misma barrera, es decir, que1 y 2 son interdependientes. Si es así, el sistema es autopoiético.

Aplicando estos criterios, se puede determinar si un sistema es autopoiético, es decir, si está vivo o no, incluso a un nivel en que la definición de la vida de sentido común ya no aporta gran cosa. En la Tabla 2 se analizan una serie de "objetos" en términos de los tres criterios arriba mencionados.

De este modo, un virus es una estructura provista de un borde, y satisface el criterio número uno, pero no así el segundo, ya que los componentes del virus no son generados al interior de la estructura del virus, sino en el exterior, en la célula huésped. Del mismo modo, un cristal que crece no está vivo, ya que los componentes del sistema no se generan desde adentro: aunque una sola molécula se replique a sí misma, no podemos decir que está viva, ya que el sistema no está dotado de una barrera semipermeable.

Observemos que a este nivel autopoiético de la definición de la vida, la evolución darwiniana no aparece como criterio de lo



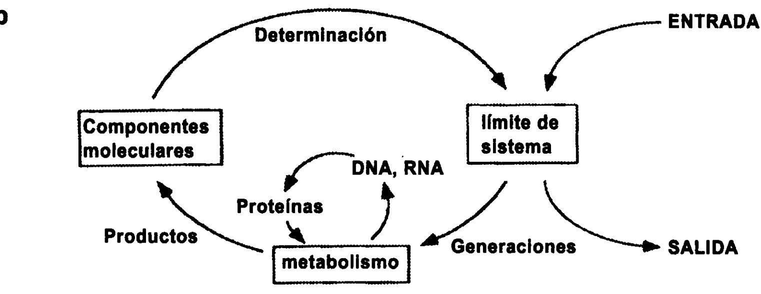


Figura 4: La circularidad del *sí mismo* en la organización autopoiética: En sus relaciones operacionales, los componentes moleculares determinan una red dinámicamente organizada y enlazada de transformación (metabolismo), que produce los componentes moleculares de la membrana, lo que, a su vez permite la operación de la red, y así sucesivamente. En nuestro mundo biológico, todo el sistema de reacciones está regulado por los ácidos nucleicos, como aparece graficado en la figura 4b. Tal como se señaló en el texto, la figura 4a representa el patrón más general, válido también para una vida «no codificada» (en la vida artificial o en los sistemas químicos sintéticos).

vivo. Esto, no porque esta propiedad no sea importante, sino simplemente porque la evolución es un hecho observable que requiere de un conjunto de unidades o entidades definidas en que el proceso de selección puede ocurrir, es decir, éste es el nivel que definimos en la introducción a nivel de la genética de una población. Es importante ahora que examinemos la relación entre los dos niveles y definiciones que mencionamos en la introducción

Objeto	criterio1: límite	criterio 2: red	criterio 3: interdependencia	conclusión
virus	sí	no	no	no
cristal	no	no	no	no
amebas	sí	sí	sí	sí
mitocondria	sí	sí	no	no
sección de ADN	no	no	no	no

Tabla 2: Vivo o no vivo, de acuerdo a los criterios de la autopoiesis

de este trabajo, para ver cómo se relacionan entre sí. Creemos que la relación es lógicamente bastante sencilla: el nivel individual de la organización autopoiética representa el patrón general de la vida mínima, válido tanto para la vida codificada y la no codificada (por ejemplo, la vida artificial) y que requiere de una población de unidades autopoiéticas caracterizada por el material genético para implementar la vida a nivel de la población genética. En otras palabras, una definición de la vida que da cuenta de la individualidad es empírica y lógicamente anterior a una secuencia evolutiva. Es totalmente posible que hayan surgido muchos sistemas vivos que no dieron origen a una descendencia viable. Por supuesto, se podría argumentar que sin evolución no hay verdadera vida, pero lo que aquí nos hemos propuesto es explorar los beneficios de establecer una separación entre el individuo y las series de tiempo. En la definición que aquí proponemos, la evolución, y la evolución darwiniana en particular, es más bien la consecuencia de la vida y no su causa.

Del mismo modo, obsérvese que el problema del origen de la vida claramente constituye el núcleo de esta dicusión sobre la definición de la vida. Pero obsérvese también que el problema de cómo y cuándo se inició realmente la lógica circular de la autopoiesis en la historia de la vida sobre la Tierra, y con qué componentes moleculares específicos, es un problema totalmente diferente.

Ahora bien, retrocedamos un poco y comparemos la definición de sentido común del extraterrestre y la definición autopoiética basada en la vida mínima. A juzgar por la representación pictórica (ver Figuras 1 y 3), hay una gran semejanza entre ambas. Sobre esta base, surge la tentación de decir que una cosa viva, macroscópica, no es más que una gigantesca unidad autopoiética, en que la envoltura externa del cuerpo físico —la piel— representa el límite. Pero no es tan simple. La autopoiesis se refiere y se limita a la vida mínima; es decir, su teoría y sus principos se basan en los organismos unicelulares. No se habla de organismos multicelulares o de jerarquías de autopoiesis. Uno podría describir una jerarquía de subsistemas autopoiéticos: por ejemplo, un órgano como el corazón o el hígado podrían ser considerados como una jerarquía de segundo orden, el animal entero como de tercer orden y así sucesivamente. Pero queremos ser rigurosos en el uso de la terminología y de las definiciones: los agregados de unidades autopoiéticas no necesariamente son autopoiéticos y, por lo demás, generalmente no lo son, así como los agregados de cuerpos humanos en una sociedad no constituyen un cuerpo humano en un orden superior. Este problema merece un análisis más profundo, pero que va más allá del propósito de este trabajo.

4. La vida a nivel de las "moléculas vivas"

La perspectiva autopoiética es esencialmente una perspectiva celular de la vida mínima. El problema que ahora hay que abordar es si la unidad celular es necesaria para una definición de la vida. Olvidémonos de la célula y observemos los mecanismos moleculares del ARN, el ADN y los ribosomas: ¿acaso no es esto suficiente para captar la esencia de la vida, para definir lo que es la vida? En otras palabras, ¿por qué no podemos definir la vida, en su expresión mínima, en términos de ciertas clases moleculares, tal como proponen algunos investigadores contemporáneos?

Esta pregunta ha cobrado gran importancia en los últimos diez años, desde el descubrimiento de los ribosomas por Cech, e independientemente, por Altman. Según parece, aquí teníamos en la misma molécula simultáneamente la capacidad de catálisis enzimático

y el contenido de información de los ácidos nucleicos. La pregunta que uno frecuentemente escucha es entonces la siguiente: supongamos que tenemos en un tubo de ensayo una molécula de ARN que es capaz de replicarse a sí misma autocatalíticamente y evolucionar por mutación: ¿no sería ésta una molecula viva?

Quisiera hacer algunas puntualizaciones antes de proceder con este análisis. Primero, hay que dejar de lado la noción de una "sola molécula" a favor de una población de moléculas. De hecho, si una sola molécula evolucionara, se transformaría, por definición, en una molécula diferente —ya no sería la molécula original. Entonces deberíamos hablar, por lo menos, de una gran familia de moléculas, para que la noción darwiniana de variación y evolución pueda aplicarse en forma más correcta. En este sentido, dejamos atrás el nivel de un solo individuo para volver al nivel de la población genética y del desarrollo histórico de la vida, el que, como señalamos al comienzo, no es el propósito de este artículo. Habiendo hecho este alcance, habrá todavía que preguntarse si algunos de los ejemplos que se propocionan en relación con el ARN no constituyen ejemplos de moléculas vivas. En particular, la definición de lo vivo, en este caso, sería: *un sistema ARN (o ADN) que es capaz de replicarse a sí mismo autocatalíticamente y de mutar en este proceso*. El punto importante que se relaciona con nuestro artículo es que en esta definición no hay ninguna noción de límite, de membrana celular, o más generalmente de una unidad emergente. Los que adhieren a esta escuela no niegan la importancia del límite, pero sólo lo consideran como un paso más que permite una autorreplicación más eficiente y no como un prerrequisito esencial.

Varios de estos sistemas han sido recientemente descritos por las escuelas de Joyce, Szostk y colaboradores, aunque no han expresado explícitamente una definición de la vida. Hay consecuencias interesantes que se desprenden de esta perspectiva acerca de la vida mínima. Por ejemplo, podría entonces también considerarse que los dinucleótidos de Rebek, que son capaces de autorreplicarse y de mutar, están vivos. Y esto nos lleva a preguntarnos si otras especies moleculares —como, por ejemplo, los oligómeros de poliestireno— que en el proceso podrían trans-

formarse en autorreproductores y capaces de mutación, por ejemplo al ser oxidados, también están vivos. Pareciera que esto plantea más problemas de los que resuelve: la definición anterior probablemente debería restringirse al ADN o al ARN —debiera ser considerada como una definición de lo vivo que es dependiente de la estructura (ADN y ARN) más que dependiente del patrón (el patrón de la autorreproducción y de la mutación). En la definición autopoiética de lo vivo ocurre todo lo contrario: lo esencial es el patrón, no la estructura. En este sentido, la definición autopoiética de lo vivo y la definición de lo vivo de la biología molecular difieren radicalmente.

Es en este punto cuando el problema de la validez de las definiciones adquiere relevancia. Como dijimos al comienzo, hay diferentes niveles desde los cuales podemos considerar/definir lo vivo: podemos buscar una definición basada en la conducta de los imputados —seres vivos vistos desde afuera— sin analizar los mecanismos que operan dentro del organismo. En ese caso, la definición de sentido común del extraterrestre resulta bastante adecuada para la vida a nivel macroscópico: y la perspectiva autopoiética constituye una buena definición del patrón más general de la vida celular mínima. Si queremos una definición de la vida mínima que haga referencia explícita a la continuidad genética y a la evolución de los procesos vitales en esta Tierra, entonces es necesario hacer referencia a los ácidos nucleicos. Obsérvese también que, por principio, cada definición debe necesariamente establecer una discriminación —la línea discriminatoria tiene que pasar por algún punto y puede que algunas personas estén disconformes con las consecuencias (por ejemplo, que se diga que un virus no está vivo): también, como toda definición, permite algunos casos fronterizos, en que la discriminación entre lo vivo y lo no vivo puede ser conflictiva.

Referencias bibliográficas

Bachmann, P.A., Luisi, P.L., y Lang, J. Autocatalytic self-replicating micelles as models for prebiotic structures. *Nature*, 357:57-59, 1992.

Bachmann, P.A., Walde, P., Luisi, P.L., y Lang, J. Self-replicating reverse micelles and chemical autopoiesis. *Journal of the American Chemical Society*, 112:8200-8201, 1990.

Bachmann, P.A., Walde, P., Luisi ,P.L., y Lang, J. Self-replicating micelles: Aqueous micelles and enzymatically driven reactions in reverse micelles. *Journal of the American Chemical Society*, 113:8204-8209, 1991.

Bernal, J.D. En T. Waterman y H.J. Morowitz, (Eds.), *Theoretical and Mathematical Biology*, 1965.

Biebricher, C.K., Eigen, M., y McCaskill, J.S. Template-directed and template-free RNA synthesis by QB replicase. *Journal of Molecular Biology*, 231:175-179, 1993.

Fleischaker, G.R. Autopoiesis: the status of its system logic. *Bio Systems*, 22:37-49, 1998.

Hargreaves, W.R. y Deamer D.W. Liposomes from ionic, single-chain amphiphiles. *Biochemistry*, 18:3759-3768, 1978.

Luisi, P.L. Defining the transition to life: self-replicating bounded structures and chemical autopoiesis. En *Thinking about Biology*. Addison-Wesley, 1993.

Luisi, P.L. y Varela, F.J. Self-replicating micelles —a chemical version of a minimal autopoietic system. *Origins of Life and Evolution of the Biosphere*, 19:663-643, 1989.

Maturana, H.R. y Varela, F.J. *Autopoiesis and Cognition —The Realization of the Living*. H.D. Reidel, Dordrecht, Holanda, 1980.

Oparin, A,L., Serebrovskaya, K,B., Pantskhava, S.N., y Vasileva, N.V. Enzymatic synthesis of polyadenylic acid in coacervate drops. *Biokhimiya*, 28(4):671-675, 1963.

Park, T.K., Feng, Q., y Rebek j, Jr. Synthetic replicators and extrabiotic chemistry. *Journal American Chemical Society*, 114:4529-4532, 1992.

Perret, M. *New Biol.*, 12:68, 1952.

Rebek, Jr. J. Molecular recognition and the development of self-replicating systems, *Experientia*, 47:1069-1104, 1991.

Stein, W. y Varela F.J. SFI studies in the science of complexity, Lect. Note Vol. III. Technical report.

Varela, F.J. *Principles of Biological Autonomy*. North Holland, Nueva York, 1979.

Varela, F.J., Maturana, H.R., y Uribe R. Organization of living systems, its characterization and a model. *Biosystems*, **5**:187-196, 1974.
von Kiedrowski, G. Ein selbstreplizierendes Hexadesoxynucleotid. *Angewandte Chemie*, **98**:932-934, 1986.
von Kiedrowski, G., Wlotzka, B., y Helbing, J. Sequenzabhängikeit matrizengesteurter Synthesen von Hexadesoxynucleotid-Derivaten mit 3′-5′-Pyrophosphatverknäpfung. *Angewandte Chemie*, **101**:1259-1261, 1989.

¿QUÉ ES LA VIDA ARTIFICIAL?

La idea y la disciplina de la vida artificial aparecieron abruptamente hacia el fin de los años 80, con promesas y nuevas problemas en la reflexión sobre los fundamentos de la vida vistos desde un modo productivo (es decir, hechos por el hombre) y no histórico. Este texto fue redactado luego de un animada discusión sobre el tema, en el Museo de Ciencias en París, donde un auditorio curioso y numeroso participó de manera decisiva.

Hay una vida "anecdótica" que tiene la virtud de golpear la imaginación, los Tamagotchis, esos pequeños perros electrónicos que uno alimenta y cuida; si son maltratados, mueren; si son bien alimentados, se reproducen. Tal vez no sean dignos de un análisis muy sofisticado pero nos demuestran que antes de plantearnos interrogantes acerca de la vida artificial, tal vez debamos preguntarnos qué es la vida ..."natural". El problema de la definición de la vida se plantea cuando uno se interroga acerca de lo que podría ser la vida extraterrestre, el nacimiento de la vida sobre la Tierra y la vida mínima. ¿Cuáles son las funciones que caracterizan lo viviente?

En el método científico habitual, las hipótesis se validan por la exactitud de las previsiones fundadas en éstas. Con la vida, el razonamiento metodológico es ligeramente diferente: se imple-

---

(Varela, F. (1997), Qu'est-ce que la vie artificielle?, *Pour la Science*, N°. 240, Octubre, pp.10-12.)
Este texto es una transcripción hecha por la revista *Pour la Science* del debate grabado, con Emile Noël, en el Palais de la Découverte, París.

mentan las cualidades hipotéticas de la vida sobre un organismo artificial y, cuando éste se comporta como un organismo vivo, la definición es validada. Lo artificial permite así juzgar las hipótesis acerca del funcionamiento de la vida. Se define, por ejemplo, lo que es una célula mínima, para fabricar un ser artificial que tiene los atributos de lo natural. Se estudian ciertas funciones del sistema neuronal, se hacen simulaciones y éstas se comparan con los verdaderos cerebros: la simulación confirma nuestra comprensión del cerebro y la veracidad de nuestra caracterización.

Este procedimiento es de uso delicado. Si lo artificial es una imitación de lo natural, ¿cuándo se transforma en prueba la imitación? Si construyo algo que que no es una imitación, pero que es cualitativamente comparable, incluso equivalente, ya no es exactamente una imitación. Sin embargo, desde nuestro punto de vista es un medio para caracterizar lo vivo. Lo vivo se caracteriza más por sus cualidades que por su modo de implementación. El tema de lo artificial, que pertenece al ingeniero de la robótica, al biotecnólogo o a las personas que se interesan en aplicaciones específicas, es una disciplina en sí. Es importante separar hoy la vertiente tecnológica y aplicada de lo artificial y la base científica fundamental. Nos interesaremos aquí en las características biológicas de la vida, en sus vínculos con las dimensiones de lo artificial ligadas a la investigación de base.

1. La compartimentación

La cuestión de la vida mínima está unida a las del origen de la vida y la vida extraterrestre. Esta impone la determinación de un criterio. ¿Cuáles eran las características de los primeros organismos vivos? ¿Qué preguntas le haremos a un extraterrestre lejano para determinar si es un ser vivo?

Dos posturas se oponen actualmente: los que piensan que la existencia del ARN, es decir los ácidos nucleicos, es una condición necesaria y suficiente, y los que estiman que se requiere algo más.

Los ARN son moléculas que tienen una doble función: son ladrillos utilizados para la construcción de estructuras complejas y tienen también una capacidad catalítica, al funcionar como enzimas. Esta doble capacidad, que no tiene el ADN, convenció a algunos de que el ARN caracterizaba lo vivo mínimo.

Sin ARN, no hay vida. No hay duda al respecto. Pero si hay ARN, ¿hay acaso vida? Mi respuesta es: no necesariamente. Es preciso que haya una frontera que separe una entidad del resto de la sopa molecular en que se encuentra.

Todos los biólogos están convencidos de que la vida no es separable de la noción de célula, pero los defensores del ARN piensan que la célula no es sino una consecuencia de las posibilidades enzimáticas del ARN. El bando del que formo parte estima indispensable la constitución de una membrana fronteriza, de un compartimento que desprende al ARN de las reacciones bioquímicas externas y que no está inscrita en el ARN.

Esta separación biológica debe ser engendrada por el elemento constitutivo de la vida mínima. De ahí la noción que elaboré hace algunos años: la *autopoiesis*, es decir la autoproducción. La vida es revelada si y solamente si la célula, en el sentido etimológico de la palabra, produce ladrillos para construir la frontera que la encierra. Hace veinte años enuncié esta idea por una serie de razones que podría desarrollar, pero desde entonces ha sido retomada por numerosos biólogos.

Para la notable bióloga Lynn Margulis, esta compartimentación autoproducida es el criterio que determina el origen de la vida, el momento en el que las macrocélulas complejas se encadenan para constituir los bordes del compartimento interno. En Marte, si queremos probar la existencia de la vida, tendremos que encontrar "células" autoproducidas, ya sea por ARN o por otras moléculas.

Podría pensarse que los biólogos carecen de imaginación en la medida en que no pueden imaginar una vida sin ARN. EL equipo de L. Luisi, en Zurich, intentó constituir células valiéndose

exclusivamente de lípidos bipolares y lo logró parcialmente, haciendo gala de una ingeniosidad experimental absolutamente admirable. Por lo tanto, la pregunta sigue abierta: tal vez hay otras dos o tres posibilidades aparte del ARN, pero las opciones son ciertamente limitadas.

2. ¿Robot experto o bebé experto?

La robótica es otra vertiente del problema. Todo el mundo vio *2001, la odisea del espacio*, donde HAL, una enorme máquina, encarna la noción de robot de los años 60-70, un aparato cognitivo extremadamente lógico, performante, con una memoria formidable, capacidades deductivas, etc. ¿Qué le falta a HAL? ¡El cuerpo! HAL no está encarnado en una flexibilidad de percepción y acción que le confiere una autonomía en relación con su mundo.

Lo que les falta a los robots no es memoria o lógica, es la simple capacidad de ver, de hacer movimientos en función de la percepción del entorno. El experto en ajedrez no es el héroe que hay que imitar, el experto es el bebé. Los bebés desprovistos de conocimiento abstracto se desenvuelven muy bien en su entorno. La parte más larga de la evolución fue el establecimiento de la capacidad de existir como una entidad autónoma, que se mueve y percibe. Los artrópodos, nuestros parientes más lejanos en la historia del planeta, no son esencialmente otra cosa. ¿Qué le falta a HAL? La facultad de moverse para hacer frente a lo inesperado del mundo. En esa medida, la simulación no es un buen test de lo natural, porque es un mundo construido sin lo inesperado.

Lo inesperado es que la vida está asociada a un mundo físico y químico en el que la variabilidad es literalmente infinita. No es agotable, y ciertas leyes pueden establecerse, pero hay una riqueza que impide la simulación y que milita en favor de la construcción.

¿Por qué lo que veo en mi computador no es algo vivo? Porque este objeto no es autónomo, no tiene existencia separada de ese universo informático inerte. En los años 50-60, se pensaba que se

podían explorar, mediante la simulación, todas las posibilidades. Pero se llegó a un callejón sin salida: el mundo es demasiado rico para que lo vivo posea una representación previa o incluso útil.

3. La reproducción es una consecuencia

No he incluido la reproducción en mi definición de la vida porque creo que sería un error epistemológico. Las huellas paleontológicas de la historia de la vida se remontan a 3,8 mil millones de años, edad de las bacterias fósiles halladas en ciertos sedimentos. Ahí, las primeras huellas detectables de vida son arqueobacterias que se parecen mucho a ciertas bacterias actuales. Sin embargo, la reproducción de las arqueobacterias no es como la de los mamíferos. En los protozoarios, la reproducción es simplemente una fisión. Si se somete un sistema autónomo autopoiético a turbulencias con olas de fluido relativamente potentes —lo que sucede con todos los sistemas acuáticos terrestres— se producen muy fácilmente fisiones.

El hecho inesperado es que la reproducción es una consecuencia banal de la existencia individual. Lo difícil es crear un individuo que se autoproduzca con su frontera. Cortarlo en dos es fácil.

Naturalmente, apenas existe un mínimo de capacidad de fisión y de división, la lógica evolutiva interviene: algunas variaciones introducen cierta diversidad y presiones selectivas que desencadenan la selección darwiniana. En unos pocos millones de años, las células han adquirido ciertas moléculas que facilitan la fisión o que favorecen fisiones selectivas con otros organismos. La reproducción es necesaria para la evolución, pero no para el origen de la vida.

4. La constitución de la identidad

¿Debe hablarse de la vida como si se tratara de una sola cosa? La vida mínima celular, es decir, la detección de un umbral de paso de lo vivo a lo no vivo, es un aspecto esencial de la vida, pero no

el único. Hay otros niveles de vida orgánicos más graduales donde se manifestan en diversos estadios las capacidades de adaptación, de evolución o incluso de adquisición de conocimientos.

En los organismos multicelulares, las unidades celulares están imbricadas en niveles de complejidad casi inverosímiles. El conjunto de células de un organismo no es una colección arbitraria, está organizado mediante comunicaciones moleculares —hormonas y, sobre todo, mediadores inmunitarios. Una identidad macromolecular se constituye para dar un cuerpo. Una frontera separa el sí mismo del no-sí mismo; ella implica, por ejemplo, el rechazo de los injertos y las reacciones febriles en cuadros de infección. Esta identidad somática difiere de la identidad cognitiva que está orquestada por el sistema nervioso fundado sobre la organización de base de percepción-acción, es decir, la pareja órganos sensoriales-músculo a través del cerebro, pero también la memoria, el aprendizaje.

Si, por desgracia, usted padece un accidente cerebral, su identidad cognitiva puede verse afectada sin que su identidad somática lo sea. O viceversa. Todos conocemos personas que tienen una actividad interna intensa, mental y espiritual, en un cuerpo reducido a nada. Son dos modos de identidad distintos que constituyen un organismo. Las dos entidades están tan bien separadas que me inclino más por el recorte en organismos que en órganos, intermediario de lo vivo demasiado tomado en cuenta en el pasado.

En el caso de los mamíferos, coexisten los sistemas somáticos y cognitivos; para las plantas, está el sistema somático y otros modos identitarios, pero que no son cognitivos, porque las plantas no se desplazan. El cerebro está esencialmente ligado a la movilidad: desde el momento en que uno se mueve, desarrolla un cerebro.

5. El jugador de ajedrez no es nada

Lo sorprendente ha sido descubrir que el secreto de la capacidad cognitiva y adaptativa no está en las lógicas, no está en lo abstracto, no está en la simple manipulación: es la capacidad de desenvolverse

con el mundo, de buscar, de orientarse, de evitar. El conocimiento no es un conjunto de procesos mentales abstractos, cuya realización más pura serían los jugadores de ajedrez.

El jugador de ajedrez me parece, por el contrario, elemental. Las reglas son fijas, en un mundo que no se mueve, y dentro de esas reglas, sólo basta con calcular. Pero la vida no está reglamentada de ese modo. Eso se comprobó en los años 60: en el MIT, se debía introducir un gran programa en una máquina que supuestamente podía resolver cualquier problema que se le planteara. La cuestión es que las preguntas por resolver no son enumerables en una lista finita. No se pueden escribir módulos de razonamiento y de acción para las diferentes funciones del cuerpo.

En 1989, un investigador del mismo instituto, Rodney Brooks, escribió un artículo que se llamaba *La inteligencia sin representación*. Preconizaba la construcción de un "bicho" que no tenía ningún saber acerca de lo que es el mundo, pero que sabía desenvolverse sobre la marcha. Este pequeño insecto podía no conocer nada acerca de las reglas de la física y de las reglas de inferencia. Estaba desprovisto de esquema esquemático o simbólico, y, sin embargo, podía ser eficaz porque estaba construido sobre la base del par percepción-acción. Inicialmente, ni siquiera tenía facultad de aprendizaje. Eso nos recuerda, naturalmente, al pequeño robot-insecto que fue enviado a Marte.

La batalla actual opone a los cognitivistas que trabajan siempre sobre un paradigma de representación de procesamiento de la información y los situacionistas que piensan que el conocimiento es una adaptación concreta, que se hace sobre la marcha y en cada situación. Para ellos, el héroe no es el jugador de ajedrez sino el bebé.

El matemático Turing había "cocinado" una definición de la inteligencia para un matemático: la máquina era inteligente cuando uno podía distinguir sus respuestas verbales de las de un ser humano. Según ese criterio, el computador *Deeper Blue* es inteligente.

Esa respuesta es poco satisfactoria para el biólogo: la inteligencia no es un conjunto de reglas, sino la capacidad de desenvolverse en un entorno cambiante.

Los objetos vivos tienen un sistema perceptivo que comanda un sistema de movimiento, el cual afecta la percepción: el animal tiene potencialmente la capacidad sensori-motriz de desenvolverse, y las soluciones para ello son sorprendentemente variadas. Observen el mundo animal donde 1000 soluciones cumplen la misma función. El lenguaje, el imaginario, son prolongaciones ulteriores, soluciones de desenvoltura.

El gran invento de la vida es la unión captor-neurona-músculo. A partir de ahí, hay una complejidad creciente de la autoorganización cerebral. Interacciones suficientemente complejas, es decir, no lineales, hacen emerger un estado global del cerebro que están en el origen de las *performances* "superiores": la imaginación, la comunicación. La investigación intenta definir las reglas de emergencia de la representación global. ¿Cómo se organizan las actividades locales de todas las neuronas en un todo comprensible? Es la primera vez en la historia de la neurobiología que los biólogos intuyen pistas para comprender cómo este enorme sistema puede funcionar de manera armoniosa y al mismo tiempo tener la capacidad de fluidez para pasar de una cosa a otra con flexibilidad y belleza.

PATRONES DE VIDA
Entrelazando identidad y cognición

Este es un intento de desplegar en detalle una idea clave en mi trabajo: la radical inseparabilidad entre la individualidad y la facultad de conocer. Esta dupla, inaugurada con el origen de la vida, se repite como tema y variaciones en todas las escalas de la evolución, hasta configurar lo que llamo aquí una "galaxia" del cierre identitario.

1. Contexto

La vocación de la Vida Artificial (VA) es contribuir a una mejor teoría científica de la vida y de los organismos vivos. Considero que la VA está guiada por dos motivaciones principales:

1. La construcción de seres/objetos vivos como medio de prueba para las posturas en litigio respecto de distintos aspectos de la vida y distintos niveles de la organización viva. Esta es la misma motivación de la inteligencia artificial con respecto a la ciencia cognitiva. Se trata de una verdadera innovación en la ciencia, ya que la física, para probar y validar, dependía principalmente de la predicción. En estos casos, podemos

---

(Varela, F. (1997), Patterns of Life: Interwining identity and cognition, *Brain Cognition* 34:72-87.)Este artículo fue concebido inicialmente para el discurso inaugural de la II[a] Conferencia Europea sobre la Vida Artificial (ECAL), Bruselas (1993). La actual versión incorpora varios puntos que surgieron durante la reunión en San Sebastián. Algunas de las ideas han aparecido además en Varela (1991). Se agradece el apoyo del CNRS.

también llegar a una validación a través de la interpretación, lo que es algo bastante diferente.

2. La utilización de los resultados de la VA para fines prácticos y tecnológicos. Creo que la circulación entre investigación e innovación tecnológica es, en gran medida, lo que ha hecho atractiva esta nueva ola de VA. Existe, nuevamente, un paralelo con la IA de comienzos de los sesenta, cuando la circulación tecnológica se estableció con sistemas expertos y otros similares. En el caso de la VA, las aplicaciones en la nanotecnología y la robótica cumplen un papel similar, al menos a los ojos del público y de los empresarios.

Está claro que la VA tiene sólidas raíces históricas y que se ha alimentado, y a la vez ha alimentado, muchas búsquedas paralelas, de las cuales es una reencarnación reciente, por así decirlo. Estas raíces comprenden el movimiento cibernético de los años cincuenta y la ciencia congnitiva y la Inteligencia Artificial de los años ochenta, pero ciertamente se trata de un árbol complejo y de muchas ramas. Con esto no pretendo negar que la VA tiene actualmente una especificidad, ampliada de manera significativa por la convergencia de los resultados anteriores y las nuevas herramientas de investigación. Ahora bien, considerando la importancia en la ciencia de saber de donde se proviene, entre otras razones para evitar la repetición de errores, las raíces históricas de la Vida Artificial deben investigarse seriamente.

También está claro que existe una cantidad de opciones epistemológicas para los diversos programas de investigación en la VA. Debemos ser claros respecto a ellos. Mi propósito aquí es bosquejar las opciones que he estado manejando en los últimos 20 años y mis razones para hacerlo.

2. El punto de vista de la autonomía

Quisiera comenzar declarando que el entendimiento de los organismos y lo viviente *es* posible, que la definición de estos términos de manera satisfactoria no es un sueño utópico, y que, incluso, tenemos una buena parte del camino recorrido. Pero, todo bajo una condición fundamental: que la autonomía de lo vivo sea resaltada en lugar de olvidada, como lo ha sido hasta ahora (Varela, 1979; Varela & Bourgine, 1991).

Destacar la autonomía significa esencialmente colocar al centro del escenario dos proposiciones interconectadas.

Proposición 1: Los organismos son fundamentalmente un proceso de constitución de una identidad.

(a) Identidad significa aquí una cualidad unitaria, una coherencia de algún tipo. No se trata de una descripción estructural estática (se trata de un proceso), ni de agregarle connotaciones mentalistas o psicológicas (se trata de una identidad en sentido general, no personalizado).

(b) La naturaleza de este proceso es siempre una de cierre operacional (Varela, 1979), esto es, un proceso de interconexión reflexiva circular, cuyo efecto primario es su propia producción.

(c) Es este cierre operacional lo que da lugar a una coherencia emergente o global, sin la necesidad de un "controlador central"; por ende, la identidad a la que aquí me refiero está localizada de una manera no substancial y, sin embargo, es perfectamente capaz de generar interacciones.

(d) Una clave esencial aquí es, por cierto, lo que hemos aprendido recientemente acerca de las propiedades "emergentes" en varios sistemas complejos.

(e) Distintos organismos difieren en los tipos de mecanismos de identidad multifarios debido a sus trayectorias evolutivas únicas.

Proposición 2: La identidad emergente del organismo proporciona, lógica y mecánicamente, el punto de referencia para un dominio de interacciones.

(a) Las identidades vivientes están producidas por algún tipo de cierre, pero lo que se produce es un nivel interactivo emergente. Las interacciones tienen relevancia y consecuencias para la identidad unitaria, aun cuando, en términos mecánicos, todas las interacciones ocurran tanto a nivel de componentes como a nivel de unidad.

(b) Las configuraciones de un nivel de interacción para la unidad completa crean una perspectiva a partir de la cual puede producirse una interacción. En otras palabras, esta es la fuente de los valores de información, intencionales o semánticos para todos los organismos vivos.

(c) Esto implica que los sistemas vivos dan lugar a significados: los organismos son autónomos, no dirigidos desde afuera, es decir, heterónomos.

Estas dos nociones clave —cómo se produce la identidad y lo que esta identidad permite en términos de la generación del dominio de existencia propio de la vida que se expresa en las Proposiciones 1 y 2— son complementarias, las dos caras de una misma moneda. Esto se ilustra en la Figura 1.

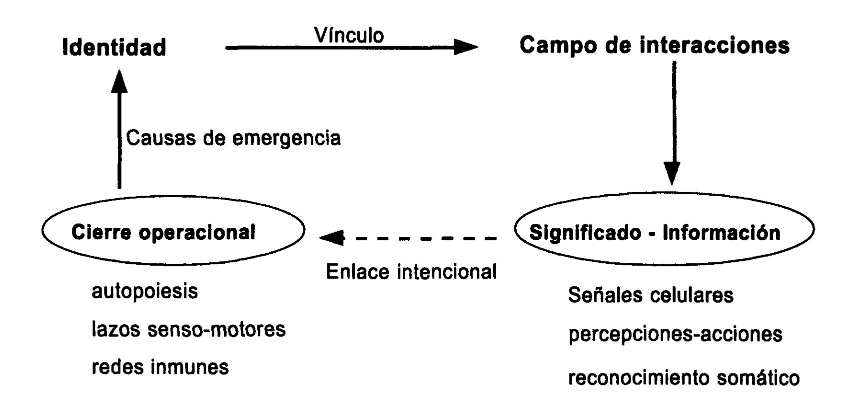


Fig. 1 Los dos conceptos clave entrelazados que dan cuenta de la autonomía de lo vivo.

Advierto al lector que mi posición es explícita, por lo que tiene la ventaja de estar abierta a la discusión. Una segunda ventaja es que podemos apuntar a dos consecuencias de esta teoría del organismo que pueden ser falsadas.

1. Ningún entendimiento adecuado de un nivel de vida será factible salvo y hasta que se haya identificado el mecanismo de identidad apropiado. Esto predice, en particular, que todos los enfoques funcionalistas están destinados al fracaso. (esto ya es bastante evidente en el estudio actual del origen de la vida).
2. El papel del acoplamiento y contingencia históricos no es secundario, sino inseparable de la existencia orgánica, ya que la identidad no es "substancial" o "abstracta". En particular, no puede haber una teoría separada de la existencia corporal, y la tarea de las simulaciones sólo puede constituir una ayuda, no una reconstitución.

3. Identidad celular

En esta sección, tomemos la estructura general expuesta en la sección 2 y articulémosla con un aspecto fundamental de los sistemas vivos: la unidad celular. Por cierto que este no es el único aspecto que necesitamos comprender, pero ilustrará en algún detalle todas las nociones claves tratadas en la Sección 2 a un nivel general.

3.1 *Autopoiesis: El proceso de la unidad celular*

La célula bacteriana constituye el sistema viviente más simple porque posee la capacidad de producir, a través de una red de procesos químicos, todos los componentes químicos que conducen a la constitución de una unidad diferenciada y limitada. Para evitar ser trivial, el atributo "viviente" en la descripción anterior debe referirse al proceso que permite dicha constitución, no a las materialidades que participan en dicho proceso ni a la enumeración de sus propiedades.

Pero ¿cuál es el proceso básico? Su descripción debe situarse a un nivel muy específico: debe ser lo suficientemente universal como para permitirnos reconocer a los sistemas vivientes como una clase, sin una referencia esencial a los componentes materiales. Al mismo tiempo, no debe ser demasiado abstracta, esto es, debe ser lo suficientemente explícita como para permitirnos ver a tales patrones dinámicos en acción en el sistema vivo que conocemos en la tierra, en aquellos que potencialmente podrían encontrarse en otro sistema solar y eventualmente en aquellos creados artificialmente por el hombre.

La biología celular contemporánea hace posible enunciar la caracterización de esta organización viva básica —una bio-lógica— como la de un sistema autopoiético (del griego: auto-producido) (Maturana & Varela, 1973, 1980). Un sistema autopoiético —la organización viva mínima— es aquel que produce continuamente los componentes que lo especifican, los cuales al mismo tiempo hacen efectivo (al sistema) como una unidad concreta en el espacio y tiempo, que a su vez hace posible la red de producción de componentes. Definido de una manera más precisa: un sistema autopoiético se organiza (se define como unidad) como una red de procesos de producción (síntesis y destrucción), de componentes, de manera que estos componentes:

(i) se regeneran continuamente y hacen efectiva la red que los produce, y

(ii) constituyen el sistema como la unidad distinguible en el dominio en el cual existen.

Esto se ilustra en la Figura 2.

Por tanto, la autopoiesis intenta capturar el mecanismo o proceso que genera la identidad de lo vivo y, así, servir como una distinción categórica entre lo vivo y lo no vivo, ni más ni menos. Esta identidad implica una coherencia autogenerada: el mecanismo autopoiético se mantendrá a sí mismo como una unidad nítida en tanto su concatenación básica de procesos se mantenga intacta

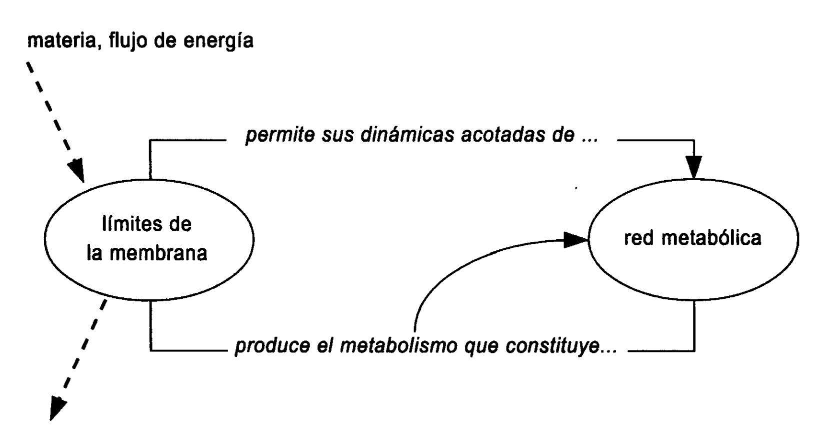


Fig 2. Esquema del cierre autopoiético de la organización viva mínima.

frente a perturbaciones y desaparecerá al ser confrontado con perturbaciones que vayan más allá de cierto rango viable, lo que depende del sistema específico que se considere. Obviamente, todos los caminos bioquímicos y la formación de membranas en las células pueden inmediatamente incorporarse a esta definición de autopoiesis.

Es importante destacar dos aspectos de la fenomenología viviente que esta caracterización autopoiética no aborda en detalle. En primero lugar, cómo esta organización autopoiética básica, presente en el origen de la vida terrestre, se hace progresivamente más compleja a través de mecanismos reproductivos, de compartimiento, dimorfismo sexual, modalidades de nutrición, simbiosis, y así sucesivamente, dando lugar a una variedad de vida pro- y eucariótica (más inteligente) en la tierra hoy en día. En segundo lugar, tomo aquí la postura de que la reproducción no es intrínseca a la lógica mínima de lo vivo. La reproducción debe ser considerada como una complejidad adicional superpuesta sobre una identidad más básica, la de una unidad autopoiética, una complejidad que es necesaria a causa de las restricciones de las condiciones primitivas en un planeta turbulento. Es en este punto donde las clases moleculares particulares juegan un papel clave, tales como los ácidos nucléicos. La reproducción es esencial para

la viabilidad de lo viviente a largo plazo, pero una unidad sólo puede reproducirse cuando existe una identidad. En este sentido, la identidad tiene una prioridad lógica y ontológica sobre la reproducción, aunque no una precedencia histórica.

En este contexto, la típica pregunta acerca de la Vida Artificial es la siguiente: ¿puede una estructura molecular más simple que la ya intrincada célula bacteriana satisfacer los criterios de la organización autopoiética? Esta pregunta puede responderse mediante la simulación y análisis de un sistema autopoiético mínimo. Los estudios de sistemas mínimos convergen en la asunción de una base de reacción de polimerización mediatizada por enzimas. Considerar, por ejemplo, (Varela, Maturana & Uribe, 1974) una trama bidimensional en la que dos tipos de elementos pueden moverse al azar: substratos (S) y unos pocos catalizadores. Una primera transformación ("composición") produce nuevos elementos como monómeros (M) que pueden enlazarse en cadenas ("enlazarse") hasta que se descomponen, con una cinética distinta a la composición. Las simulaciones muestran la emergencia espontánea de una cadena de polímeros que se enlaza consigo misma. En esta simulación permitimos una permeabilidad diferencial a través de la cadena de polímeros (por ejemplo, nula para el catalizador, muy baja para los monómeros, alta para el substrato) y una aproximación al movimiento de Brown. Curiosamente, estas unidades autodistintivas no sólo surgen a través de estas sencillas normas, sino que su enlazamiento está dotado de cierta autorregulación, como cuando el deterioro se ha producido en un segmento de membrana. Dados los rangos diferenciales de difusión y deterioro, el mismo límite de la unidad permite su propia reparación y vuelta a un estado de unidad, nuevamente distinto de su entorno.

Así, la simulación ilustra la emergencia de una unidad dotada de límites, que involucra la naturaleza paralela y distribuida de procesos similares a los procesos químicos. Quisiera agregar que usamos este autómata celular una década antes que se hicieran populares como casos ideales para estudiar las propiedades emergentes. La simulación antes indicada es un intento explícito de producir un sistema autopoiético mínimo y, en este sentido, es

bien diferente del propósito de otros autómatas celulares, que buscan propiedades distintas de la autodistinción, tales como la reproducción (en el caso del juego de la vida de Conway) o patrones espacio-temporales (en el caso de las clasificaciones de Wolfram).

¿Qué pasa con estos sistemas autopoiéticos mínimos en el medio químico actual, y en cuanto a su relevancia para el origen de la vida? De hecho, el encapsulamiento de las macromoléculas por vesículas lípidas ha sido activamente investigado como un candidato prometedor para una célula primitiva (ver Deamer, 1985). Luisi & Varela (1989) sostienen que un sistema micelular reversible podría estar muy cerca de ser un sistema autopoiético mínimo. Recientemente, el grupo de Luisi ha tenido éxito en producir sistemas micelulares que contienen en su núcleo acuoso una reacción conducente a la producción de un surfactante, el cual constituye una barrera para la reacción micelular reversible (Bachman, Luisi & Lang, 1992; Luisi, 1993). Lo interesante de este caso es que se sabe mucho acerca de estos sistemas químicos que hacen efectivamente posible poner en funcionamiento un sistema autopoiético mínimo.

Hasta ahora me he referido al tema del organismo como un sistema viviente mínimo a través de la caracterización de su modalidad básica de identidad. Esto de acuerdo con la Proposición I antes indicada, lo que en rigor implica abordar el tema a un nivel ontológico: el énfasis está puesto sobre la manera en que un sistema viviente se convierte en una entidad distinguible y no en su composición molecular específica y sus configuraciones históricas contingentes. En tanto exista, la organización autopoiética permanece invariable. En otras palabras, una manera de resaltar la especificidad de la autopoiesis es pensar en su autorreferencia como aquella organización que mantiene a la organización misma como algo que no varía. Toda la constitución físico-química está en constante flujo; el patrón permanece, y solamente a través de esta invariabilidad organizativa puede determinarse el flujo de los componentes. En particular, si bien los sistemas autopoiéticos son sistemas químicos disipativos, una pura caracterización materia-energía de ellos desconoce por completo la arquitectura específica o circuitos de materiales (por ejemplo, la Figura 2) de estos sistemas

químicos altamente complejos que inauguran el dominio de los individuos autónomos, y por ende de la vida misma.

Insisto: me he referido aquí solamente a la organización celular mínima que da lugar a la vida autónoma. Como dije, mi propósito es destacar la bio-lógica básica que sirve como fundamento a partir del cual puede considerarse la diversidad visible en los organismos actuales: solamente cuando existe una identidad pueden verse las elaboraciones como familias de variaciones de una clase común de unidades vivas. Cada clase de entidad tiene una identidad que le es peculiar; el carácter único de lo vivo reside en el tipo de organización que tiene.

Ahora bien, la historia de la biología está, por cierto, desfigurada por la oposición tradicional entre los mecanicistas/reduccionistas por un lado y los holistas/vitalistas por otro, una herencia del problema-espacio biológico del siglo XIX. Una de las contribuciones específicas del estudio de los mecanismos autoorganizativos —de los cuales la autopoiesis es una instancia específica— es que la oposición tradicional entre los elementos componentes y las propiedades globales desaparece. En el simple ejemplo del autómata celular ilustrado arriba, es precisamente la causalidad recíproca entre las normas locales de interacción (es decir, las normas de los componentes, las cuales son similares a interacciones químicas) y las propiedades globales de la entidad (es decir, su demarcación topológica, la cual afecta la difusión y crea las condiciones locales para la reacción) lo que queda en evidencia. Me parece a mí que esta causalidad recíproca contribuye mucho a terminar con la oposición mecanicista/vitalista y nos permite avanzar a una fase más productiva de la identificación de diversas modalidades de autoorganización en las cuales lo local y lo global aparecen entrelazados explícitamente a través de esta causalidad recíproca. La autopoiesis es un ejemplo claro de la dialéctica entre los niveles locales de componentes y el todo global, enlazados en una relación recíproca a través de la exigencia de la constitución de una entidad que se separa de su trasfondo. En este sentido, la autopoiesis entendida como la caracterización de

los patrones básicos de lo vivo no cae en los extremos tradicionales ni del vitalismo ni del reduccionismo.

3.2 La identidad de lo vivo y su mundo

La segunda dimensión complementaria de la bio-lógica básica, central para enfocar nuestra discusión, es la naturaleza de la relación entre unidades autopoiéticas autónomas y su entorno, en conformidad con la Proposición 2. Independientemente de la hipótesis, resulta evidente que un sistema autopoiético depende de su entorno físico-químico para su conservación como una entidad autónoma, de lo contrario se disolvería y reintegraría a su entorno. De aquí la intrigante paradoja propia de una identidad autónoma: el sistema vivo debe diferenciarse de su medio ambiente y al mismo tiempo debe mantener su vinculación con él; este vínculo no puede deshacerse ya que el organismo emerge desde el medio ambiente pero a la vez se debe al mismo.

Ahora bien, en este enlace dialógico entre la unidad viva y el entorno físico-químico existe una diferencia clave del lado de lo vivo, dado que cumple un papel activo en este enlazamiento recíproco. Al definir lo que constituye una unidad, en el mismo gesto se define lo que queda fuera de dicha unidad, esto es, el entorno circundante. Un examen más detenido hace también evidente el hecho de que esta organización sólo puede ser entendida, por decirlo así, desde el "interior": la unidad autopoiética crea una perspectiva desde la cual tiene un exterior que le es propio, que no puede ser confundido con el entorno físico nuestro en tanto observadores, el ámbito de las leyes físicas y químicas.

En nuestra práctica de biólogos, nos movemos en uno y otro terreno. Usamos y manipulamos principios y propiedades físico-químicos, pero de inmediato podemos utilizar interpretaciones y significados vistos desde el punto de vista del sistema de vida. Así, una bacteria que nada en un gradiente de sucrosa es analizada por conveniencia en términos de los efectos locales de la sucrosa en la permeabilidad de la membrana, viscosidad del medio,

hidromecanismo del batir de los flagelos, etc. Pero por otra parte, el gradiente de sucrosa y el batir de los flagelos son dignos de ser estudiados sólo en la medida en que son relevantes para todas las bacterias: su significado específico como componentes de conducta alimenticia es posible sólo por la presencia y perspectiva de las bacterias como una totalidad. Si consideramos a la bacteria como una unidad, todas las correlaciones entre los gradientes y las propiedades hidrodinámicas se transforman en leyes químicas ambientales, evidentes a nuestros ojos de observadores, pero desprovistas de toda significación.

Creo que esta relación claramente dialéctica es de suma importancia. Incluso, puede parecer tan obvia que no alcancemos a apreciar sus profundas ramificaciones. Me refiero a la distinción entre el medioambiente de un sistema de vida tal como se presenta al observador y sin referencia a una unidad autónoma —lo que de ahora en adelante llamaremos simplemente el medioambiente— y el medioambiente para el sistema, que se define en el mismo movimiento que dio lugar a su identidad y que sólo puede existir en esa definición mutua —a partir de ahora, el mundo del sistema.

La diferencia entre medioambiente y mundo es el excedente de significación que acosa al entendimiento de la vida y del conocimiento y, a la vez, está en la base de cómo un sí mismo alcanza su individualidad. En la práctica es bastante difícil no perder de vista la dialéctica de esta definición mutua: ni aislamiento rígido ni sencilla continuidad con la química física. Por el contrario, es fácil combinar el mundo de las unidades con su medioambiente dado que es obvio que estamos estudiando ésta o aquella interacción molecular en el contexto de una unidad celular autónoma y, en consecuencia, perder completamente el excedente agregado por la perspectiva del organismo. No hay significación alimenticia en la sucrosa, salvo cuando una bacteria remonta el gradiente y su metabolismo utiliza la molécula de modo tal que permite la continuidad de su identidad. Este excedente no es indiferente a las regularidades y las texturas (por ejemplo "las leyes físicas") que operan en el medioambiente, a que la sucrosa pueda crear un

gradiente y atravesar la membrana de una célula, etc. Opuestamente, el mundo del sistema esta construido sobre estas regularidades, lo que asegura que pueda mantener su unidad en toda circunstancia.

Lo que hace el sistema autopoiético —a causa de su particular modo de identidad— es enfrentar constantemente los encuentros (perturbaciones, golpes, acoplamientos) con su medioambiente y tratarlos desde una perspectiva que no es intrínseca a los encuentros mismos. Es seguro que las rocas o las cuentas de cristal no atraen gradientes importantes de azúcar de sus infinitas posibilidades de interacciones físico-químicas, porque para que esto suceda es esencial una perspectiva desde una identidad constituida activamente. Es tentador, en este punto, dejarse resbalar hacia las vaporosas nubes del "significado", reminiscencias de la peor clase de vitalismo del pasado o de la jerga informática del presente. Lo que quiero hacer notar aquí es que aquello que es significativo para un organismo, está dado precisamente por su constitución como proceso distributivo, con una indisoluble unión entre los procesos locales en los que ocurren interacciones (por ejemplo, las fuerzas físico-químicas actuantes en una célula) y la entidad coordinada que equivale a la unidad autopoiética, dando lugar al manejo de su medioambiente sin necesidad de acudir a un agente central que mueva los controles desde afuera (como un *élan vital*) o un orden preexistente en una localización particular, como un programa genético que espera ser expresado.

Si invertimos nuestra perspectiva, esta constante producción de significado puede ser descrita como una permanente falta en lo vivo: está constantemente produciendo una significación que está perdida, nunca pre-dada o pre-existente. La relevancia debe ser adquirida *ex nihilo*: distinguir especies moleculares relevantes e irrelevantes, seguir un gradiente hacia arriba y no hacia abajo, aumentar la permeabilidad de este ion y no de aquel, etc. Se manifiesta un inevitable *contratiempo* entre un sistema autónomo y su medioambiente: siempre hay algo que el sistema debe proveer desde su perspectiva de totalidad en función. En el hecho, un encuentro molecular adquiere significación en el contexto del

sistema operacional completo y de muchas interacciones simultáneas.

La fuente de esta creación-de-mundos es siempre un quiebre en la autopoiesis, ya sea menor como los cambios en la concentración de un metabolizado, o mayor como la ruptura de los bordes. Debido a la naturaleza misma de la autopoiesis —ilustrada por la reparación de la membrana en el ejemplo mínimo simulado más arriba—, todo quiebre puede verse como el inicio de una acción, por parte del sistema sobre lo que falta, de manera que se mantenga la identidad. Repito: no hay una teleología implícita en este "de manera que": esto es lo que implica la lógica autorreferencial de la autopoiesis, en primer lugar. La acción se hará visible como un intento de modificar su mundo: cambio de lugar de diferentes nutrientes, aumento de la corriente de un metabolizado por síntesis metabólica, etc.

En resumen, esta acción permanente e inexorable sobre lo que falta, se convierte, desde el punto de vista del observador, en la actividad cognitiva del sistema, que es la base de la diferencia inconmensurable entre el medioambiente en cuyo interior es observado el sistema, y el mundo, en cuyo interior opera el sistema. Esta actividad cognitiva se encuentra paradójicamente en su base misma. Por una parte, la acción que produce un mundo es un intento por reestablecer una relación con un medioambiente que desafía la coherencia interna mediante encuentros y perturbaciones. Pero estas acciones, al mismo tiempo, demarcan y separan al sistema de ese medioambiente, dando lugar a un mundo distinto.

El lector puede sorprenderse del uso que hago del término cognitivo, hablando de sistemas celulares. Sin embargo, de lo que he dicho hasta ahora debería quedar claro que la constitución de un dominio cognitivo vincula a los organismos y a sus mundos de una manera que constituye la esencia misma de la intencionalidad como se la usa en la ciencia cognitiva moderna, y como fuera introducida originalmente en la fenomenología. Mi proposición hace explícito al proceso a través del cual surge esta intencionalidad: implica una hipótesis explícita acerca de cómo

transformar esta noción filosófica de intencionalidad en un principio para las ciencias naturales. El uso que hago aquí del término cognitivo se justifica porque está en la base misma de como la intencionalidad surge en la naturaleza.

En resumen, el término cognitivo tiene dos dimensiones constitutivas: primero, su dimensión de enlace, esto quiere decir, la conexión con su medioambiente que le posibilita mantener su individualidad; segundo su dimensión interpretativa, esto quiere decir, el excedente de significación que adquiere una interacción física, a causa de la perspectiva otorgada por la acción global del organismo.

4. Identidad cognitiva

4.1 *La percepción-acción como una neuro-lógica básica*

En la sección anterior, he presentado el entrelazamiento fundamental entre identidad y cognición, las Proposiciones 1 y 2, tal como opera en el caso de un organismo mínimo. En esta sección, quisiera mostrar cómo el nivel más tradicional de propiedades cognitivas, incluyendo los cerebros de los animales multicelulares, es, en una importante medida, la continuación del mismo proceso básico.

El cambio desde una celularidad mínima a un organismo con sistema nervioso es rápido, y pasa por alto la complejidad de las múltiples maneras que tienen los organismos multicelulares para surgir y evolucionar. Esta es una transición que sucede en unidades de selección, e implica el balance somático de poblaciones diferenciadas de células en un organismo adulto, así como el esforzado desarrollo de caminos hacia el establecimiento de una estructura corpórea. Tal como Buss ha afirmado recientemente: "La evolución del desarrollo es la generación de una 'ecología somática', que mediatiza los posibles conflictos entre célula e individuo, mientras que el organismo interactúa de manera simultánea y efectiva con el entorno extrasomático" (Buss, 1987, página 12). Para efectos de lo que me interesa aquí, asumiré con fines de utilidad la identidad de un organismo multicelular, distinto de una entidad

autopoiética mínima en su modalidad de identidad, pero similar en el sentido que demarca a una entidad autónoma de su entorno (Varela & Frenk, 1987).

Ahora bien, ¿cuál es el lugar específico del sistema nervioso en la operación corporal de un multicelular? Toda vez que el movimiento es una parte integral del estilo de vida de un multicelular, hay un desarrollo del sistema nervioso que relaciona los efectores (músculos, secreción) y las superficies sensoras (órganos sensoriales, terminaciones nerviosas). La lógica fundamental del sistema nervioso es la relación de los movimientos con una corrientes de modulaciones sensoriales, de manera circular. La red resultante son correlaciones de percepción y acción que surgen y son moduladas por un conjunto de neuronas activas, la red de interneuronas. Consecuentemente, las neuronas son únicas entre las células de un organismo multicelular en sus ramificaciones axonales y dentríticas, que permiten contactos múltiples que se extienden a grandes distancias (relacionadas al tamaño del soma celular) proveyendo el medio esencial para esta correlación entre los efecto-sensores y el intra-organismo.

Contrariamente a la forma habitual en que se describen estas arquitecturas neuronales, desearía poner énfasis desde el comienzo el carácter "en-situación" de esta neuro-lógica: el estado de actividad de los sensores resulta típicamente de los movimientos del organismo. En gran medida, la conducta es la regulación de la percepción. Esto no excluye, por supuesto, perturbaciones independientes del entorno. Pero lo que generalmente se describe como "estímulo" en el laboratorio, una perturbación que es deliberadamente independiente de la actividad que está realizando el animal, es menos pertinente (fuera del laboratorio), para entender la biología de la cognición.

Permítanme volver atrás un momento y reformular nuestra discusión sobre el ser cognitivo junto a aquello del ser molecular mínimo. Postulo que la neurociencia contemporánea —al igual que la célula biológica para el caso de la organización de la vida— da suficientes elementos para la concepción de la organización básica

de un ser cognitivo en términos de cierre operacional (no interaccional) del sistema nervioso (Maturana & Varela, 1980; Varela, 1979). Me refiero a cierre para destacar la cualidad autorreferencial de las redes de interneuronas y de las superficies percepto-motoras cuya correlación favorece. La calificación de operacional enfatiza que el cierre es usado en su sentido matemático de recursividad, y no en el sentido de encierro o aislamiento de la interacción, lo cual sería un absurdo, por supuesto. En términos más específicos, el sistema nervioso está organizado mediante operaciones de cierre de una red de sub-redes modular y recíprocamente relacionadas, dando origen a conjuntos de actividad coherente tales como:

i) Patrones de correlación sensorimotora que son estables;

ii) Una conducta para todo el organismo como unidad móvil en el espacio. Esto queda ilustrado en la Figura 3.

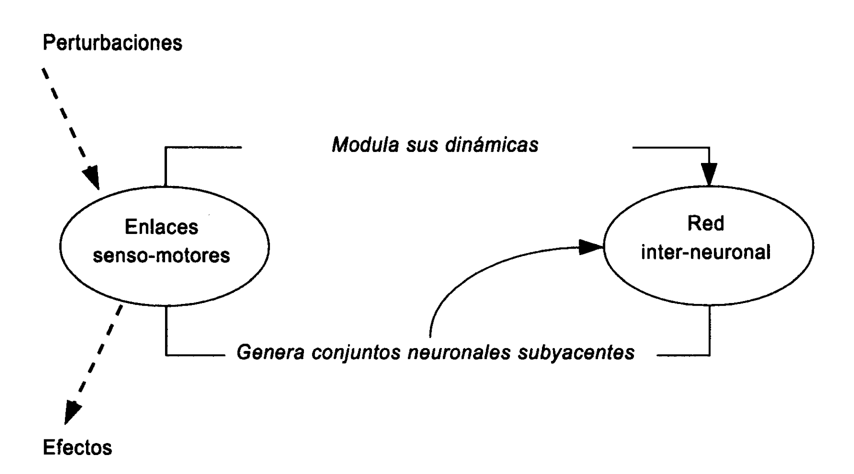


Fig. 3. Esbozo del cierre operacional del sistema nervioso.

El cierre operacional del sistema nervioso produce un modo específico de coherencia, inscrustado en el organismo. Esta coherencia es un identidad cognitiva: una unidad de percepción/movimiento en el espacio, de constantes senso-motoras que son medidas a través de la red de interneuronas. El paso a la cognición ocurre en el nivel de la entidad conductual, y no, como en el ser celular básico, como una entidad espacialmente limitada. La clave en este proceso cognitivo es el sistema nervioso través de su neuro-lógica. En otras palabras, el ser cognitivo es la manera en que el organismo, mediante su actividad autoproducida, llega a ser una entidad distinta en el espacio, aunque siempre acoplada a su correspondiente medioambiente, del cual, no obstante, será siempre diferente. Un ser distinto coherente que, por el proceso mismo de constituirse (a sí mismo), configura un mundo externo de percepción y acción.

4.2 *Ser cognitivo y mundo perceptual*

La naturaleza de la identidad neuro-cognitiva que acabamos de discutir, como la del ser celular básico, es de emergencia a través de un proceso distributivo. Lo que deseo traer a colación aquí es la relativamente reciente (¡y aplastante!) conclusión de que muchos agentes simples con propiedades simples pueden reunirse, incluso al azar, para dar nacimiento a lo que aparece ante el observador como un todo múltiple e integrado, sin necesidad de una supervisión central. Ya hemos tocado este tema cuando discutimos la naturaleza del proceso autopoiético y el modelaje autómata celular, y más adelante cuando discutimos el constante auge y caída del conjunto neuronal que está en el origen del comportamiento. Este punto es crucial para toda mi argumentación: he basado mis conclusiones en estudios contemporáneos sobre una variedad de sistemas complejos inspirados en la biología, que abarcan casi todo el campo de trabajo de los investigadores sobre IA. Introduce una alternativa explícita a la tradición computacional dominante en el estudio de las propiedades cognitivas, donde la idea central es el procesamiento de información externa sucesivamente elaborada

para reconstituir una representación central. Este paradigma fundamental del programa computacional no sirve para la biología, ni para la IA.

Me he detenido en este punto para advertir al lector sobre la fuerza acumulada en muchos años de dominación del lenguaje computacional, y su consecuente tendencia a identificar el ser cognitivo con algunos programas o descripciones computacionales de alto nivel. Esto no sirve. El ser cognitivo es su propia implementación: su historia y su acción son una sola cosa. Ahora bien, esto exige que aclaremos el segundo aspecto del ser que nos interesa tratar: su modo de relacionarse con el medioambiente.

La vida cotidiana es necesariamente una vida de agentes situados, que llevan adelante sus quehaceres enfrentados a una constante de actividades paralelas en su variados sistemas perceptuomotores. La continua redefinición de los quehaceres no es de ninguna manera un plan, guardado en un repertorio de alternativas potenciales, sino una realidad dependiente de la contingencia, de la improvisación, y más flexible que planificada. Estar situado significa que la entidad cognitiva tiene, por definición, una perspectiva. Esto significa que no está "objetivamente" relacionada con su medioambiente, que es independiente del sistema de locación, títulos y apartados, actitudes e historia. En cambio, se relaciona con el medioambiente desde la perspectiva establecida por la constante emergencia de características del agente mismo y en términos del papel que juegan tales redefiniciones en la coherencia del sistema.

Un vez más, tal como hicimos con el ser celular mínimo, debemos hacer una diferencia tajante entre medioambiente y mundo. Y una vez más la manera de acoplamiento es doble. Por una parte, este cuerpo-en-el-espacio claramente sucede a través de las interacciones con el entorno del cual depende. Estas interacciones tienen el carácter de encuentros macrofísicos —transducciones sensoriales, fuerza y rendimiento muscular, luz y radiaciones, etc.—, nada de qué sorprenderse. Sin embargo, este acoplamiento sólo es posible si los encuentros son asumidos desde la perspectiva del sistema mismo.

Esto equivale, rigurosamente, a elaborar una significación extrarrelativa a esta perspectiva. Todo lo que se encuentre deberá ser valorado de una u otra manera —gusto, disgusto, ignorancia— y recibido de una u otra manera —atracción, rechazo, neutralidad. Esta afirmación básica es inseparable de la manera en que el acto de acoplamiento se encuentra con una unidad perceptuo-motora en funcionamiento, y da lugar a una intención, esa cualidad única de la cognición viva.

Dicho en otros términos, la naturaleza del medioambiente adquiere para un ser cognitivo un curioso status: es la que se presta a un excedente de significación. Al igual que la improvisación en jazz, el medioambiente otorga la "excusa" para la "música" neural desde la perspectiva del sistema cognitivo involucrado. Al mismo tiempo, el organismo no puede vivir sin este acoplamiento constante y la igualmente constante emergencia de regularidades: de faltar la posibilidad de actividad asociaciativa el sistema se convertiría en un simple fantasma.

Por ejemplo, la luz y el reflejo (entre otros parámetros macrofísicos como bordes y texturas, pero simplifiquemos para argumentar mejor), se prestan a una gran variedad de gama de colores, dependiendo del sistema nervioso involucrado en tal encuentro. Durante sus respectivos caminos evolutivos, los peces teleósteos, pájaros, mamíferos e insectos han generado diferentes gamas de colores, no solamente con distintas significaciones conductuales, sino, además, con diferentes dimensiones, de manera que no se trata de mayor o menor resolución de los colores (Thompson, Palacios, & Varela, 1992). Otro ejemplo patente de este excedente de significación y la sorprendente gestión del cerebro como generador de "narrativas " neuronales se puede encontrar en la tecnología de la realidad virtual. Lo que me parece más significativo, en este caso, es la veracidad del mundo creado: nos sentimos habitantes de un cuerpo en este nuevo mundo después de breve lapso de prueba (más o menos 15 minutos) y la experiencia es de un verdadero vuelo o de desintegración en universos fractales. Esto es así a pesar de la mala calidad de la imagen, la baja sensibilidad de los sensores y la limitada

interconexión entre las superficies sensorial y de imagen, a través de un programa instalado en un computador personal. En su completud, el sistema nervioso es un sintetizador de regularidades tan magnífico, que cualquier material básico le es suficiente como medioambiente para producir un mundo extraordinario.

Esta misma estrategia de estar-en-situación de un agente, que es progresivamente provisto de módulos cada vez más ricos de autoorganización interna, se está convirtiendo en un productivo programa de investigación, incluso en un campo tan pragmáticamente orientado como el de la robótica, como queda bien ilustrado en los trabajos de R. Brooks o L. Steels. Es interesante observar cómo Brooks rastrea también el origen de lo que describe como "la decepción de la IA" por la tendencia a la abstracción (no solamente en la IA, sino también en el resto de las ciencias cognitivas), por ejemplo, en la fabricación de percepciones situadas y habilidades motoras. Tal como he sostenido aquí (y como discute Brooks por sus propias razones), tal abstracción desconoce la esencia de la inteligencia cognitiva, que es encontrable sólo en su inscripción corporal o encarnación. Es como si uno pudiera separar los problemas cognitivos en dos partes: una que puede ser resuelta mediante la abstracción y otra que no. La segunda incluye las capacidades motoras y de percepción-acción de agentes situados en medioambientes no especificados. Cuando nos acercamos al problema desde esta perspectiva autosituada, no hay un lugar donde la percepción pudiera entregar una representación del mundo a la manera tradicional. El mundo se muestra mediante la representación de las regularidades perceptuo-motoras. "Así como no hay representación central tampoco hay sistema central. Cada rango de actividad conecta directamente percepción y acción. Es el observador de la criatura quien atribuye una representación central o un control igualmente central. La criatura en sí misma no tiene ninguno: es una colección de conductas competitivas. Del caos local producido por sus interacciones, surge, a los ojos del observador, un modelo de conducta coherente" (Brooks, 1987, p. 11).

Para concluir, los dos puntos que he intentado aclarar en esta sección dedicada al ser cognitivo son los siguientes: primero, he

tratado de describir la naturaleza de su identidad como un cuerpo en movimiento-y-espacio, a través del cierre operacional de la red interneuronal. Esta actividad es observable como múltiples sub-redes, actuando en paralelo y entretejidas en complejos *bricolages*, generando de manera sucesiva e incesante modelos coherentes que se manifiestan como comportamientos. Segundo, he intentado aclarar cómo esta dinámica emergente, paralela y distributiva, es inseparable de la constitución del mundo, que no es otra cosa que el excedente de significado y de intenciones que aparece en los comportamientos situados. Si las conexiones con el ambiente físico son inevitables, lo que hace único al ser cognitivo es esta génesis constante de significado. O, invirtiendo la descripción, lo que hace único al ser cognitivo es su falta constitutiva de significación, que debe ser resuelta en el enfrentamiento permanente con las perturbaciones y rupturas propias de la vida perceptuo-motora. La cognición es acción referida a aquello que falta, visto desde la perspectiva de un ser cognitivo que siente aquella falta.

5. La galaxia del cierre

Las dos últimas secciones ilustran la estrategia propuesta en la sección 2 para programas de investigación sobre IA, que centran la atención en la autonomía de lo vivo. No es necesario agregar que estoy seguro de la existencia de una amplia constelación de trabajos que siguen de modo general, cuando no específico, este tipo de marco teórico. Esto queda ilustrado en la figura 4, donde "la galaxia del cierre" aparece sazonada con diversas contribuciones en varios niveles de identidad y varios modos de análisis desde lo empírico a la simulación.

Es también evidente, a esta altura, que pongo marcado énfasis en el análisis explícito de las bases epistemológicas que orientan mis investigaciones sobre el conocer natural y artificial. Me parece útil terminar, por lo tanto, contrastando mi idea de la IA basada en la autonomía, con otras distinguidas tendencias en el mismo campo de estudio. Distingo por lo menos tres paradigmas opuestos.

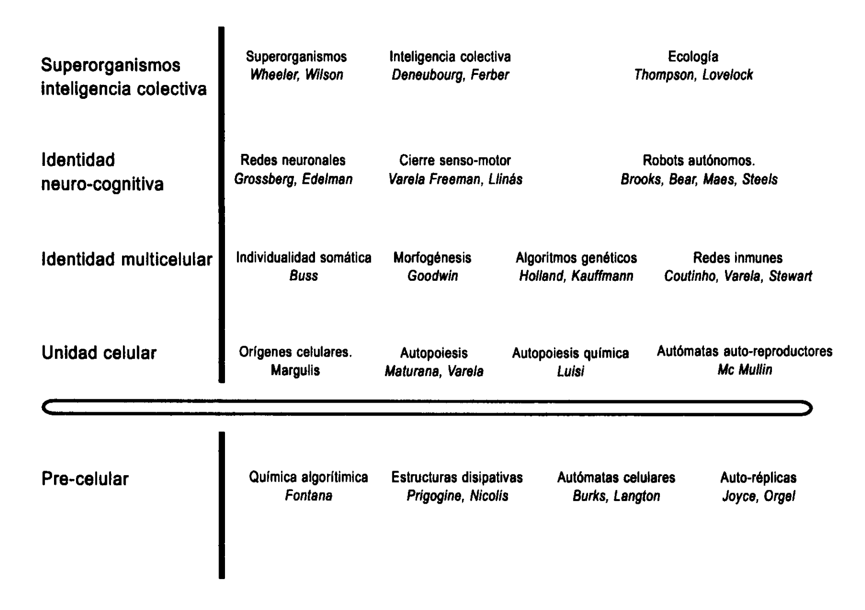


Fig. 4. Panorama de la "galaxia del cierre" con autores cuyo trabajo toca varios niveles de identidad y mecanismos específicos de cierre.

1. *Funcionalismo.* Su estrategia central se basa en una duplicación de la función o propiedad como indicador de adecuación. Dicho en breve: mientras veamos, en un sistema artificial, que una propiedad reaparece, es que estamos avanzando hacia alguna parte y podemos seguir confiando en la simulación. El funcionalismo ha alcanzado popularidad mediante los trabajos de investigadores como C. Langton o D. Dennet en las ciencias cognitivas, y de manera más descarada a través de las obras de escritores populares como S. Levy.

2. *Fisicalismo.* El punto central en esta teoría es la confianza en ciertas descripciones físico-moleculares, contrastadas como algunas propiedades emergentes explícitas. Recientemente, R. Penrose se ha convertido en un férreo defensor de esta postura al utilizar la mecánica cuántica para explicar la conciencia.

3. *Anti-naturalistas.* Aquí la idea central es que la tarea de definir científicamente la vida, el organismo y el conocimiento, está destinada al completo fracaso. Sus representantes más tradicionales son vitalistas de varias clases, pero es bueno reconocer que hoy en día van quedando pocos. Los más comunes siguen siendo dualistas estrictos, como J. Eccles o, algo más interesantes, neomisterianistas, como T. Nagel.

Una discusión detallada de los distintos paradigmas haría necesario otro artículo. Ofrezco este mero esbozo para dejar claro, hasta donde sea posible, cual es mi propia posición en esta ecología diversa de escuelas sobre el conocer. El futuro dirá quién estaba en el mejor camino. De cualquier modo, me parece que estamos embarcados en algo y que desde ya tenemos los cimientos de una buena teoría sobre la vida y el organismo.

Referencias bibliográficas

Bachman, P. A., Luisi, P.L. & Lang, J. 1992. Autocatalytic self-replicating micelles as models for prebiotic structures, *Nature*, 357, 57-59.

Brooks, R.A. 1987. Intelligence without representation, *Artificial intelligence laboratory report.* Cambridge, MA: MIT.

Buss, L. 1987. *The evolution of individuality.* Princeton: Princeton Univ. Press.

Deamer, D. W. 1985. Role of amphilic compounds in the evolution of membrane structure on the early Earth, *Origins of Life.* 17, 3-25.

Luisi, P.L. 1993. Chemical models of minimal autopoietic systems. In W. Stein & F. Varela (eds.) *Thinking about biology*, Nueva York: Addison-Wesley.

Luisi, L & Varela, F. 1989. Self replicating micelles: A minimal version of a chemical autopoietic system. *Origins of Life*, 19.

Maturana, H. & Varela, F. 1973. *De máquinas y seres vivos: Una teoría de la organización biológica.* Santiago de Chile: Editorial Universitaria.

Maturana, H. & Varela, F. 1980. *Autopoiesis and cognition: The realization of the living*. Boston: D. Reidel.

Thompson, E., Palacios, A. & Varela, F. 1992. Ways of coloring: Comparative color vision as a case study in cognitive science. *Behavior and Brain Science*, 15, 1-75.

Varela, F. 1979. *Principles of biologial autonomy*. Nueva York: North-Holland/Elsevier.

Varela, F. 1991. Organism: A meshwork of selfless seleves. In A. Tauber (Ed.), *Organism and the origins of self*. Dordrecht: Kluwer Academic. Pp. 79-107.

Varela F. & Bourgine, P. Eds. 1991. *Towards a practice of autonomous systems*. Cambridge: MIT Press.

Varela, F. & Frenk S. 1987. The organ of form: Towards a biological theory of shape. *Journal of Social Biology an Structure*, **10**, 73-83.

Varela, F., Maturana, H. & Uribe R. 1974. Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, 5, 187-195.

I

ORGANISMO

2. Organismo

EL ORGANISMO, UNA TRAMA DE IDENTIDADES SIN CENTRO

Este texto, más que ningún otro, desarrolla lo que yo entiendo por organismo en su sentido más propio. La noción clave es la de identidades sin centro, quizá la revolución conceptual más importante en ciencias de las últimas décadas. Puede también ser leído como otra proyección o corte del texto anterior, que pone el acento sobre la identidad (lo que explica la repetición de ciertas ideas). Aquí el acento está puesto sobre la curiosa naturaleza de las identidades que la vida ha generado.

1. Introducción

La idea de organismo implica una dialéctica complicada: un sistema viviente se estructura a sí mismo como una entidad distinta a su medioambiente mediante un proceso que genera, a través de este mismo proceso, un mundo adecuado para él.

Mi intención en lo que sigue es profundizar en esta afirmación, tanto en el sentido de otorgarle una justificación de hecho, biológica, así como de revelar algunas de sus consecuencias epistemológicas. Utilizo el término dialéctica para describir propiedades que se relacionan de manera que "... una cosa no puede existir sin la otra, que una adquiere sus propiedades de su relación con la otra, que las propiedades de ambas se desarrollan como consecuencia de sus interacciones" (1,pp. 2,3). Pero hay mucho más en esto, como veremos. En realidad, mi conclusión será que la relación entre organismo

(Varela, F., (1991), Organism: A meshwork of selfless selves, in: Tauber (ed.), *Organism and the Origin of Self*, Kluwer Assoc, Dordrecht, pp.79-107).

y el sí mismo (*self*) resulta ser la imbricación de *dos* dialécticas distintas: una ligada al mecanismo de la identidad, la otra unida al modo de relacionarse con el mundo.

La temática del sí mismo surge de inmediato: comencé diciendo que un sistema vivo se estructura a sí mismo como una entidad distinta. Muchos biólogos preferirían evitar la noción de "sí mismo". Porque ¿qué quiere decir exactamente? La tentación es pedir disculpas y comenzar a dar explicaciones del tipo: "Bueno, se trata de una especie de, ejem, no una cosa sino una actividad... o algo como eso". Otros se referirán a la distinción entre el sí mismo y aquello que no lo es, por razones inmunológicas. Incluso unos pocos se referirán al "sí mismo" como una subjetividad, el fantasmal "yo" que arrastramos con nosotros y al que prestamos nuestro nombre.

La posición básica que adopto en este trabajo es que este tema no se resolverá con respuestas vagas: una acabada cuenta acerca de la constitución de un sí mismo autónomo y de la naturaleza de su modo de existencia, está en el centro mismo de una investigación biológica y cognitiva. Necesitamos proveernos de instrumentos más poderosos para ver cómo un sí mismo puede, al mismo tiempo, ser un punto virtual sin coordenadas localizadas, y entregar una identidad a través de la cual se puede producir una interacción. Desde mi punto de vista, nociones recientes respecto de propiedades emergentes en sistemas modulares, altamentes distribuidos, hacen posible ciertos lineamientos generales para caracterizar a estos seres virtuales.

Otro planteamiento básico —en oposición a cierta exigencia de parquedad malentendida— es que la significación múltiple del concepto de "sí mismo" se entiende mejor si se lo enfoca como totalidad y no de manera fragmentada. En otra palabras: para discutir lo que podría significar el sí mismo es esencial que nos situemos en el área de todas aquellas situaciones en donde el *para* sí mismo (*pour-soi*) aparece. Al mismo tiempo cada una de las modalidades del sí mismo tiene sus propias características, como trataré de puntualizar más adelante. Necesitaremos entonces

entendernos con una multiplicidad de sí mismos regionales, todos ellos poseedores de un modo de autoconstitución y capaces, a través de su interacción y montaje general, de dar origen a un organismo. Consecuentemente, deseo invocar aquí los siguientes sí mismos "regionales": 1) una unidad mínima o celular; 2) un ser corpóreo en su base inmunológica; 3) un sí mismo perceptuo-motor asociado a la conducta animal; 4) un "yo" socio-lingüístico de subjetividad y 5) el colectivo social compuesto por la totalidad de las multi-individualidades. En todas estas áreas estamos tratando con niveles y procesos en donde emerge una identidad —no como sustancia sino como movimiento—, cuya fábrica de articulaciones es el organismo. Hacer desaparecer la multiplicidad de esta red es una fuente de confusión.

El lector no debe preocuparse: no estoy a punto de lanzarme en un viaje por todas estas áreas mediante rápidas pinceladas. Lo que intentaré, en las siguientes dos secciones (2 y 3), es discutir las regiones una y tres con suficiente detalle como para llegar a algunas conclusiones fundamentales acerca del organismo y el sí mismo en conjunto, las que serán presentadas en la sección 4.

2. El sí mismo como sistema vivo mínimo

2.1 *Autopoiesis como esqueleto bio-lógico*

La célula bacterial es el sistema más simple de vida en tanto posee la capacidad de producir, a través de una red de procesos químicos, todos los elementos químicos que conducen a la constitución de una unidad distinta y conectada. El atributo de "viviente" de la descripción anterior, debe referirse al proceso que permite tal constitución, no a la materialidad que lo compone, o a una enumeración de propiedades. Pero ¿en qué consiste este proceso básico? Su descripción debe situarse en un nivel muy específico: debe ser lo suficientemente universal como para permitirnos reconocer sistemas de vida como una clase, sin referencias esenciales a los elementos materiales que lo componen. Pero, al mismo tiempo, no debe ser demasiado abstracta,

es decir, debe ser lo suficientemente explícita como para permitirnos ver los patrones dinámicos en acción en el actual sistema de vida que conocemos sobre la tierra, aquellos que potencialmente existen en otros sistemas solares y eventualmente aquellos creados artificialmente por el hombre. Tal como lo dijera recientemente el organizador de un encuentro sobre vida artificial: "Solamente cuando seamos capaces de ver *la vida como la conocemos* en el amplio contexto de *la vida como podría ser*, estaremos preparados para entender realmente la naturaleza de la bestia" (2, p.2).

La biología celular contemporánea posibilita caracterizar esta organización vital básica —una bio-lógica—, como un sistema *autopoiético* (del griego: autocreado) (3,4,5). Un sistema autopoiético —la organización vital mínima—, es aquel que produce continuamente los componentes que lo especifican, al mismo tiempo que construye el sistema como una unidad concreta en espacio y tiempo, lo cual hace que la red de producción de componentes sea posible. Una definición más precisa es la siguiente: un sistema autopoiético esta organizado (definido como unidad) como una red de procesos de producción (síntesis y destrucción) de componentes tales como los siguientes:

(i) Regenera y construye continuamente la red que los produce y

(ii) constituye el sistema como una unidad distinguible en el dominio en el cual existe.

Por tanto, la autopoiesis intenta capturar el mecanismo o proceso que genera la identidad de la vida, y servir, por lo tanto, como una distinción categórica entre la vida y la no vida (6). Esta identidad equivale a una coherencia autoproducida: el mecanismo autopoiético se mantendrá como una unidad distinta mientras la concatenación básica de los procesos permanezca intacta ante las perturbaciones, y va a desaparecer cuando se confronte con perturbaciones que vayan más allá de cierto rango de viabilidad, lo cual dependerá del sistema específico que se considere. Obviamente, todos los caminos de la bioquímica y de la formación de las membranas en las células pueden ser inmediatamente reconsiderados de acuerdo a esta definición de autopoiesis.

Un ejercicio diferente —que no pretendo hacer ahora—, es ver cómo esta organización autopoiética básica, presente en el origen de la vida terrestre (6), va paulatinamente haciéndose más compleja a través de mecanismos reproductivos, de compartimentación, dimorfismo sexual, modos de nutrición, simbiosis, etc., dando origen a la variedad de vida pro- y eucariótica sobre la tierra (7,8). En especial, considero aquí la idea de que la reproducción *no* es

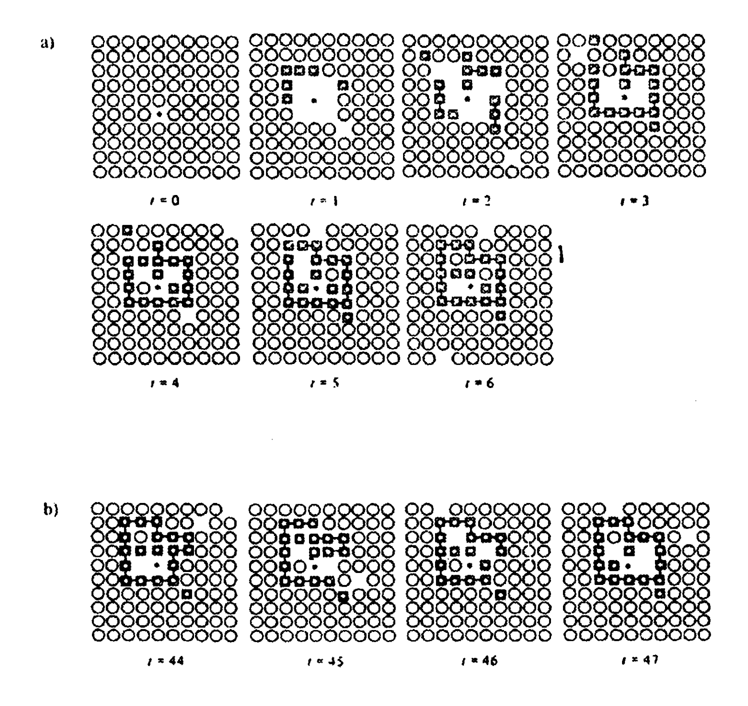


Fig. 1 Simulación de un sistema autopoiético mínimo (5). (a) Los primeros siete instantes de un sistema de "reacciones" simulado en el computador y descrito en el texto, en los cuales se muestra el surgimiento espontáneo de la unidad ligada en este espacio bidimensional. Las interacciones entre el sustrato 'O' y el catalizador '*' producen una cadena de eslabones conectados 'U' que finalmente encierran al catalizador, delimitando, por tanto, un interior para las reacciones del sustrato. (b) Cuatro instantes sucesivos en la misma corrida (a), en los cuales se muestra la regeneración espontánea de los límites, rotos por el debilitamiento de los eslabones ligados. La producción que prosigue al interior de la unidad hace posible el restablecimiento de los límites de la unidad a la manera de cambios en la forma geométrica y movimiento de los componentes.

intrínseca a la lógica mínima de la vida. La reproducción debe ser considerada como una complejidad agregada y sobreimpuesta a una identidad más básica antes que a una unidad autopoiética, una complejidad que necesariamente se debe a las restricciones producto de las duras condiciones en los primeros tiempos de un planeta. La reproducción es esencial para la viabilidad de la vida, pero sólo cuando hay una identidad se puede reproducir una unidad. En este sentido, la identidad tiene prioridad lógica y ontológica sobre la reproducción, aun cuando no tenga precedencia histórica.

Pero en lugar de seguir estas complejidades históricas, me gustaría plantear una cuestión más pertinente para mis propósitos: ¿puede una estructura molecular más simple que la ya intrincada célula bacterial, satisfacer el criterio de una organización autopoiética? Esta pregunta puede responderse mediante la simulación y el análisis del sistema autopoiético *mínimo*. Todos los sistemas mínimos parecen converger en la suposición de un sistema de polimerización mediado por enzimas. Consideremos por ejemplo (5) un cuadriculado bidimensional donde dos clases de elementos pueden moverse al azar: el sustrato ('O') y unos pocos catalizadores ('*'). Una primera transformación ('composición') produce nuevos elementos tales como monómeros, los cuales pueden unirse en cadenas ('vinculación'), hasta que decaen, con una kinética diferente a la de la composición.

Resumiendo:

composición:	* + 2 [O] —> * + O
vinculación:	[O]-[O]-[O]-... + [O] —> [O]-[O]-[O]-[O]-...
decaimiento:	[O] —> 2 O

Para visualizar lo que estas normas pueden generar, la figura 1a muestra la secuencia de una simulación, con el surgimiento espontáneo de una cadena de polímeros que giran sobre sí mismos. En esta simulación permitimos una permeabilidad diferencial a través de la cadena de polímeros (por ejemplo, cero para el catalizador, muy

baja para los monómeros, alta para el sustrato), y una aproximación al movimiento browniano.

Es interesante observar que estas unidades autodistintivas no solamente emergen mediante estas reglas sencillas, sino que, además, su capacidad de girar sobre sí mismas está dotada de cierta autorregulación. Esto se puede ver en la figura 1b, donde ocurre un decaimiento en un segmento de membrana. Dadas las jerarquías de difusión diferencial y decaimiento, el mismo límite de la unidad permite su propia reparación en una unidad, otra vez distinta a sus antecedentes.

Por lo tanto, la simulación es una instancia de unidad *emergente*, que incluye, en paralelo, la naturaleza distributiva de los procesos de tipo químico, como sucede en otros muchos ejemplos, como el tan conocido Juego de la Vida, de Conway (9), o en las clases de autómatas celulares examinadas recientemente en detalle, (10,11). La simulación presentada más arriba es un intento explícito de producir un sistema autopoiético mínimo, y en este sentido es claramente distinta de la mencionada célula autómata, la cual busca propiedades distintas a la autodistinción, tales como la reproducción (en el caso de Conway) o patrones espacio-temporales (en el caso de Wolfram). El punto crucial de las propiedades emergentes ha sido aquí tocado; volveré sobre ello más adelante.

¿Qué pasa con tales sistemas autopoiéticos mínimos en la química actual y, en consecuencia, con su relevancia para el origen de la vida? En el hecho, la encapsulación de las macromoléculas mediante vesículas lípidas ha sido investigada intensamente, como una promisoria candidata para un primera célula. Luisi y Varela (16) han propuesto por su parte que un sistema micelular reversible puede estar cerca de ser un sistema autopoiético mínimo. Específicamente, ellos discuten el caso de un sistema micelular reversible que alberga en su centro acuoso una reacción que conduce a la producción de un surfactante, que opera como límite para la reacción micelular reversible. Lo interesante es que se sabe mucho acerca de este sistema químico, lo que hace posible que hoy se pueda poner en operación un sistema autopoiético mínimo.

2.2 *La identidad de la vida y su mundo*

He tratado el tema del organismo como un sistema de vida mínimo, caracterizando su *modo básico de identidad*. Esto es, en rigor, tratar el tema en un nivel ontológico: el acento está puesto en la manera en que un sistema de vida llega a ser una entidad distinguible, y no en su composición molecular específica y las configuraciones históricas contingentes. Mientras exista, la organización autopoiética permanecerá invariable. En otras palabras, una manera de entender la especificidad de la autopoiesis es pensar en su autorreferencia como una organización que mantiene la organización misma con carácter de invariable. Toda su constitución físico-química está en constante flujo: pero los patrones permanecen y sólo a través de su invariabilidad se puede determinar el flujo de los componentes.

Me he referido aquí solamente a la organización mínima que da origen a tal autonomía de vida. Tal como he dicho, mi propósito es destacar la base biológica que sirve de fundamento desde el cual puede considerarse la diversidad visible de los organismos actuales. Sólo cuando hay una identidad, las elaboraciones pueden ser vistas como familias de variaciones de una clase común de unidades de vida. Toda clase de entidades tiene una identidad que le es particular: lo único de la vida reside en la *clase* de organización que posee.

Ahora bien, la historia de la biología está marcada por la tradicional oposición entre mecanicismo/reduccionismo por una parte y el holismo/vitalismo por la otra, como herencia del problema-espacio biológico del siglo XIX. Una de las contribuciones específicas del estudio de los mecanismos autoorganizados —de los cuales la autopoiesis es una instancia específica—, es que la tradicional oposición entre los elementos componentes y las propiedades globales desaparece. En el simple ejemplo de automatón celular ilustrado antes, es precisamente la *causalidad recíproca* entre las leyes locales de interacción (p.e. las reglas de los componentes, que son análogas a las interacciones químicas) y las propiedades globales de la entidad (su demarcación topológica que afecta la difusión y crea las condiciones locales para la reacción) la que se hace evidente. Me

parece que esta causalidad recíproca es importante para desechar la oposición mecanicista/vitalista, y, a la vez, nos permite avanzar hacia una fase más productiva en la identificación de *varios modos* de autoorganización, en donde lo local y lo global están juntos y explícitamente entretejidos a través de esta causalidad. La autopoiesis es el primer ejemplo de la dialéctica entre los niveles de componente locales y el todo global, unidos en una relación recíproca a través de las exigencias de constitución de una entidad que se autosepara de sus antecedentes. En este sentido, la autopoiesis, como la caracterización de la vida, no cae en los extremos tradicionales del vitalismo o el reduccionismo.

Una segunda dimensión complementaria de bio-lógica básica, central para focalizar nuestra discusión, es la naturaleza de *la relación* entre unidades autónomas autopoiéticas y sus ambientes. Es *teóricamente* evidente que un sistema autopoiético depende de su ambiente físico-químico para su conservación como entidad separada, de otro modo se disolvería en dicho ambiente. He aquí la intrigante paradoja de una identidad autónoma: el sistema de vida debe distinguirse de su medio, mientras que *al mismo tiempo* debe mantenerse ligado a él: esta unión no puede deshacerse por cuanto el organismo proviene precisamente de dicho ambiente. Ahora bien, en esta unión dialógica de la unidad de vida y el ambiente físico-químico, el equilibrio está ligeramente cargado hacia la vida, dado que tiene el rol activo. Al definir lo que es en tanto unidad, en ese mismo momento define lo que queda fuera de ella, es decir, el ambiente que la rodea. Un examen más exhaustivo hace también evidente que esta exteriorización sólo puede ser entendida, por así decirlo, desde "adentro": la unidad autopoiética *crea una perspectiva* desde la cual el exterior es uno, que no puede confundirse con el alrededor físico tal como aparece ante nosotros en tanto observadores; el campo de las leyes físicas y químicas *en sí mismas*, libre de la perspectiva del observador.

En nuestra práctica de biólogos, nos movemos en uno y otro terreno. Usamos y manipulamos principios y propiedades físico-químicos, pero de inmediato podemos utilizar *interpretaciones* y significados vistos *desde* el punto de vista del sistema de vida.

Entonces, una bacteria que nada en un gradiente de sucrosa es analizada por conveniencia en términos de los efectos locales de la sucrosa en la permeabilidad de la membrana, viscosidad del medio, hidromecanismo del batir de los flagelos, etc. Pero, por otra parte, el gradiente de sucrosa y el batir de los flagelos, son dignos de ser estudiados sólo en la medida en que son relevantes para todas las bacterias: su significado específico como componentes de conducta alimenticia es posible sólo por la presencia y perspectiva de las bacterias como una totalidad. Si consideramos a la bacteria como una unidad, todas las correlaciones entre los gradientes y las propiedades hidrodinámicas se transforman en leyes químicas ambientales, evidentes a nuestros ojos de observadores, pero desprovistas de toda significación.

He dado todas estas explicaciones porque creo que esta relación realmente dialéctica es de suma importancia. En realidad puede parecer tan obvia que no alcancemos a apreciar sus profundas ramificaciones. Quiero decir la distinción entre el medioambiente *de* un sistema de vida tal como se presenta al observador y sin referencia a una unidad autónoma —lo que de ahora en adelante llamaremos simplemente el *medioambiente*— y el medioambiente *para* el sistema, que se define en el mismo movimiento que dio lugar a su identidad y que sólo puede existir en esa definición mutua —a partir de ahora *el mundo* del sistema.

La diferencia entre medioambiente y mundo es el excedente de *significación* que acosa al entendimiento de la vida y del conocimiento y, a la vez, está en la base de cómo un sí mismo alcanza su individualidad. En la práctica es bastante difícil no perder de vista la dialéctica de esta definición mutua: ni aislamiento rígido ni sencilla continuidad con la química física. Por el contrario, es fácil combinar el mundo de las unidades con su medioambiente dado que es obvio que estamos estudiando esta o aquella interacción molecular en el *contexto* de una unidad celular autónoma y, en consecuencia, perder completamente el excedente *agregado* por la perspectiva del organismo. No hay significación alimenticia en la sucrosa, salvo cuando una bacteria remonta el gradiente y su metabolismo utiliza la molécula de modo tal que permite la

continuidad de su identidad. Este excedente no es indiferente a las regularidades y las texturas (por ejemplo "las leyes") que operan en el medioambiente, a que la sucrosa pueda crear un gradiente y atravesar la membrana de una célula, etc. Opuestamente, el mundo del sistema está construido *sobre* estas regularidades, lo que asegura que pueda mantener su unidad en toda circunstancia.

Lo que hace el sistema autopoiético —debido a su particular modo de identidad— es enfrentar constantemente los tropiezos (perturbaciones, golpes, acoplamientos) con su medioambiente y tratarlos desde una perspectiva que no es intrínseca a los tropiezos mismos. Es seguro que las rocas o las cuentas de cristal no atraen gradientes importantes de azúcar de sus infinitas posibilidades de interacciones físico-químicas, porque para que esto suceda es esencial una perspectiva *desde* una identidad constituida activamente. Es tentador, en este punto, dejarse deslizar hacia las vaporosas nubes del *significado*, reminiscencias de la peor clase de vitalismo del pasado o de la jerga informática del presente. Lo que quiero remarcar aquí es que aquello que es significativo para un organismo, está dado precisamente por su constitución como proceso distributivo, con una indisoluble unión entre los procesos locales en los que ocurren interacciones (por ejemplo, las fuerzas físico-químicas actuantes en una célula) y la entidad coordinada que equivale a la unidad autopoiética, dando lugar al manejo de su medioambiente sin necesidad de acudir a un agente central que mueva los controles desde afuera (como un *élan vital*) o un orden preexistente en una localización particular, como un programa genético que espera ser expresado.

Si invertimos nuestra perspectiva, esta constante producción de significado puede ser descrita como una permanente *falta* en lo vivo: está constantemente produciendo una significación que está *perdida*, nunca pre-dada o pre-existente. La relevancia debe ser adquirida *ex nihilo*: distinguir especies moleculares relevantes e irrelevantes, seguir un gradiente hacia arriba y no hacia abajo, aumentar la permeabilidad de este ion y no de aquel, etc. Se manifiesta un inevitable contratiempo entre un sistema autónomo y su medioambiente: siempre hay algo que el sistema debe proveer desde su perspectiva de totalidad en función. En el hecho, un en-

cuentro molecular adquiere significación en el contexto del sistema operacional completo y de muchas interacciones simultáneas.

La fuente de esta creación-de-mundos es siempre "un quiebre" en la autopoiesis, ya sea menor como los cambios en la concentración de un metabolizado, o mayor como la ruptura de los bordes. Debido a la naturaleza misma de la autopoiesis —ilustrada por la reparación de la membrana en el ejemplo mínimo simulado más arriba—, todo colapso puede verse como el inicio de una acción, por parte del sistema sobre lo que falta, de manera que se mantenga la identidad. Repito: no hay una teleología implícita en este "de manera que": esto es lo que implica la lógica autorreferencial de la autopoiesis, en primer lugar. La acción se hará visible como un intento de modificar su mundo: cambio de lugar de diferentes nutrientes, aumento de la corriente de un metabolizado por síntesis metabólica, etc.

En resumen, esta acción permanente e inexorable sobre lo que falta, se convierte, desde el punto de vista del observador, en la actividad *cognitiva* del sistema, que es la base de la diferencia inconmensurable entre el medioambiente, en cuyo interior es observado el sistema, y el mundo, en cuyo interior opera el sistema. Esta actividad cognitiva se encuentra paradójicamente en su base misma. Por una parte, la acción que produce un mundo es un intento por reestablecer una relación con un medioambiente que desafía la coherencia interna mediante encuentros y perturbaciones. Pero estas acciones, al mismo tiempo, demarcan y separan al sistema de ese medioambiente, dando lugar a un mundo distinto.

El lector puede sorprenderse del uso que hago del término "cognitivo" hablando de sistemas celulares. Como dije antes, uno de los puntos que me interesa hacer notar aquí es que ganamos al observar la *continuidad* entre este nivel fundamental del sí mismo y otros sí mismos regionales, incluyendo el neural y el lingüístico, de modo que no debemos titubear en el uso de la palabra "cognitivo". Supongo que otros preferirían introducir aquí la palabra "información". Sí, pero hay razones por las cuales creo que esto sería incluso más problemático. Aun cuando sea claro que podemos describir un factor X que perturba desde el exterior del organismo,

X no es información. En rigor, para el organismo sólo se trata de un *eso*, de un *algo*, un material básico para in-formar desde su propia perspectiva. En términos físicos hay material, pero no está dirigido a nadie. En el momento en que hay cuerpo —incluso en su forma más mínima— se transforma en algo in-formado para un sí mismo, en la dialéctica recíproca que expliqué recién. Esta in-formación no es nunca una significación fantasmal o bit de información a la espera de ser recolectado por el sistema. Es una presentación, una ocasión para un enlace, y es, en este *entre-deux*, que surge la significación (17, 18, 19).

Por lo tanto, el término cognitivo tiene dos dimensiones constitutivas: primero su dimensión de *enlace*, esto quiere decir, la conexión con su medioambiente que le posibilita mantener su individualidad; segundo —admito que forzando un tanto el lenguaje— su dimensión *imaginaria*, esto quiere decir, el excedente de significación que adquiere una interacción física, debido a la perspectiva otorgada por la acción global del organismo. Muchos lectores no estarán de acuerdo conmigo y considerarán el uso que hago de "cognitivo" e "imaginario" como abusos de lenguaje. A estos lectores les ruego que sean indulgentes y que me acompañen hasta el final de mi argumentación, para comprobar los méritos o deméritos de este tipo de análisis.

3. El sí mismo como sistema cognitivo básico

3.1 *La percepción-acción como una neuro-lógica básica*

En la sección anterior, he presentado el entrelazamiento fundamental entre identidad y cognición, tal como opera en el caso de un organismo mínimo. En esta Sección quiero mostrar cómo el nivel más tradicional de propiedades cognitivas, incluyendo el cerebro de los animales multicelulares, es en cierto sentido *la continuación* del mismo proceso básico.

El cambio desde una celularidad mínima a un organismo con sistema nervioso es rápido, y pasa por alto la complejidad de

las múltiples maneras que tienen los organismos multicelulares para surgir y evolucionar (20, 21, 22). Esta es una transición que sucede en unidades de selección, e implica el balance somático de poblaciones diferenciadas de células en un organismo adulto, así como el esforzado desarrollo de caminos hacia el establecimiento de una estructura corpórea. Tal como Buss ha afirmado recientemente: "La evolución del desarrollo es la generación de una 'ecología somática', que mediatiza los posibles conflictos entre célula e individuo, mientras que el organismo interactúa de manera simultánea y efectiva con el entorno extrasomático" (21).

Para la mayoría de los vertebrados, esta "ecología somática" está unida por una red de linfocitos que constituyen el núcleo del sistema inmunológico. Nuevamente, no es mi propósito iniciar aquí una discusión sobre un "sí mismo inmunológico" (*inmunological self*). No puedo resistir la tentación, sin embargo, de señalar que en otros trabajos he presentado *in extenso* un acercamiento al sistema inmunológico y su papel en la fijación de un ser celular/molecular flexible durante la ontogenia de los mamíferos (23). En mi opinión, esta identidad no es, tal como se ha afirmado tradicionalmente, una delimitación del "sí mismo" como defensa *en contra* del "no- sí mismo" compuesto de antígenos invasores (24, 25). Es una afirmación autorrefente, de una unidad coherente, —una "ecología somática"—, mediatizada a través de inmunoglobulinas libres y marcadores celulares en intercambio dinámico (26, 27). Las reacciones de inmunidad ante las infecciones, aun cuando importantes, están mediatizadas por un sistema inmunológico "periférico", una subpoblación de linfocitos movilizados, no a través de una red, sino por medio de mecanismos de expansión clonal, como una reactividad refleja adquirida a través de la evolución (28, 29). Pero basta de *excursus*. Para mi propósito aquí voy a asumir de inmediato la identidad de un organismo multicelular, claramente diferente de una entidad autopoética mínima en su modo de identidad, pero similar en cuanto a que delinea una identidad autónoma con respecto a su entorno (30).

Ahora bien, ¿cuál es el lugar específico del sistema nervioso en la operación corporal de un multicelular? Toda vez que *el movimiento*

es una parte integral del estilo de vida de un multicelular, hay un desarrollo del sistema nervioso que relaciona los efectores (músculos, secreción) y las superficies sensoras (órgano sensoriales, terminaciones nerviosas). *La lógica fundamental del sistema nervioso es la relación de los movimientos con una corriente de modulaciones sensoriales, de manera circular.* La red resultante son correlaciones de percepción y acción que surgen y son moduladas por un conjunto de neuronas activas, la red de *interneuronas*. Consecuentemente, las neuronas son únicas entre las células de un organismo multicelular en sus ramificaciones axonales y dentriticas, que permiten contactos múltiples que se extienden a grandes distancias (relacionadas al tamaño del soma celular) proveyendo el medio esencial para esta correlación entre los efecto-sensores y el intra-organismo.

Contrariamente a lo habitual, desearía poner énfasis desde el comienzo en la *puesta en situación* de esta neuro-lógica: el estado de actividad de los sensores resulta típicamente de los movimientos del organismo. En gran medidad, la conducta es la regulación de la percepción. Esto no excluye, por supuesto, perturbaciones independientes del entorno. Pero lo que generalmente se describe como "estímulo" en el laboratorio, una perturbación que es deliberadamente independiente de la actividad que está realizando el animal, es menos pertinente (fuera del laboratorio), para entender la biología de la cognición.

Las coherencias perceptuo-motoras que describimos externamente como conductas, disimulan el surgimiento, al interior de la red interneuronal, de un gran subgrupo —lo que se llama un *ensemble*— de neuronas correlacionadas transientemente. Estos conjuntos son, a la vez, la fuente y el resultado de la actividad de los sensores y las superficies efectoras. Consideremos, por ejemplo, la *Aplysia*, un molusco de agua con un sistema nervioso "pequeño" (unos pocos miles de neuronas) que ha sido intensamente estudiado (31). Cuando la *Aplysia* pone en contacto su sifón con una superficie (o el sifón es tocado separadamente), contrae su branquia. Esta es la llamada reacción "repliegue branquial", uno de los muchos patrones de conducta presentes en estos animales. Es un estudio reciente (32) se utilizó tintura sensible a voltaje para controlar una

parte importante de las neuronas de los ganglios abdominales del molusco *Aplysia.* Contrariamente a la imagen libresca de la conducta como un camino arco-reflejo, los autores concluyeron que al menos 300 neuronas están activas durante esta acción, lo cual representa una proporción significativa de la red total de la red interneuronal. El conjunto de neuronas observadas a través de este método surge en coordinación e influencia mutua, y su co-activación se debilita en pocos segundos.

Es por supuesto el caso que, aquellas interneuronas y neuronas motoras participantes en esta especial conducta de repliegue branquial, también participan en otros conjuntos que tienen que ver con otras conductas. Por lo tanto, las interneuronas de este ganglio invertebrado deben concebirse como una red de conjuntos superpuestos que se presentan en varias configuraciones coherentes, dependiendo del contexto donde el animal se encuentra (fig. 2).

La lección de este humilde molusco se extiende al resto de los animales con cerebro. Lo que cambia es la cantidad de interneuronas mediadoras y la estructura específica del respectivo sistema nervioso, con varias regiones corticales, capas y núcleos. En los humanos, alrededor de 10^{11} interneuronas interconectan algo así como 10^{6} neuronas motoras, relacionadas con 10^{7}neuronas sensoras, distribuidas en las superficies receptoras a través de todo el cuerpo. Esta es la relación de 10: 100.000: 1 de interneuronas que median el acoplamiento de las superficies sensoras y motoras. El auge y caída de la autoorganización neuronal ilustrada en el caso del retiro del sifón del molusco *Aplysia*, es más que válido en cerebros mayores. Así, por ejemplo John *et. al.* (33) descubrió que de 5 a 100 millones de neuronas se activan en el cerebro de un gato durante la simple tarea motora-visual de agarrar un trozo de hígado. Estos conjuntos de neuronas surgen en un entramado de regiones, demostrando al parelelismo distributivo enorme, propio del cerebro de los vertebrados.

En rigor, es justo decir que un hecho recientemente establecido sobre la constitución del cerebro es *La Ley de Reciprocidad*: si una región, (digamos un área cortical, o un núcleo específico) A conecta con otra región B, luego B conecta recíprocamente con A, aunque

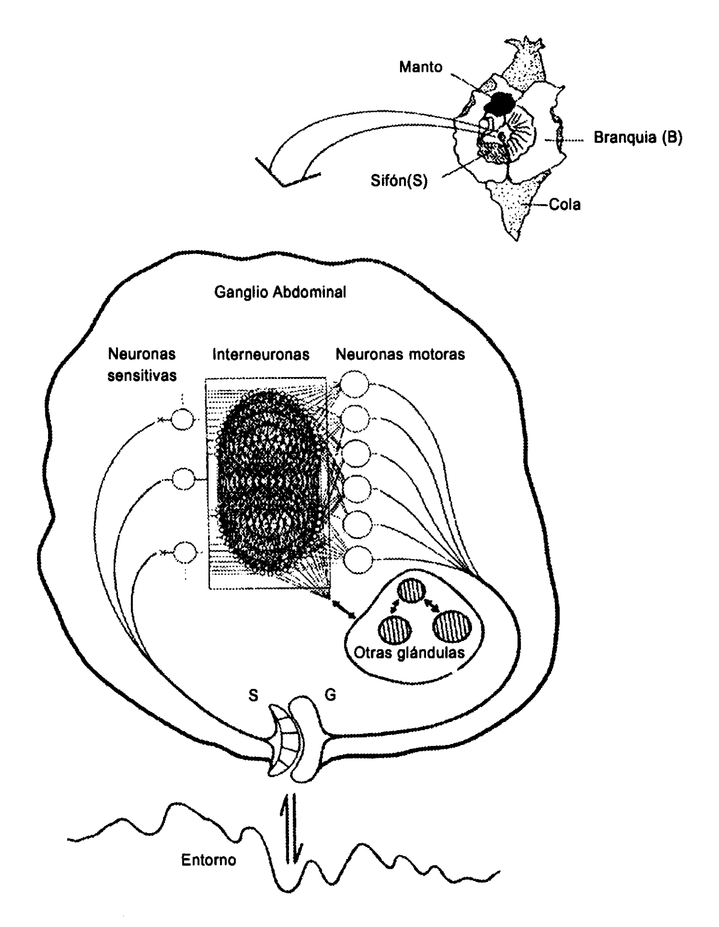


Fig. 2 Esquema de la organización básica del ganglio abdominal de la *Aplysia*, acorde con el descubrimiento de que varios cientos de neuronas son activadas cuando el sifón viene a tocar levemente un borde. La superficie sensitiva indicada corresponde a células mecanosensitivas con terminales en la piel del sifón (S). Las motoneuronas, en el diagrama, son para indicar esos múltiples músculos inervados que regulan la posición de la branquia (B). Reciben directamente conexiones (monosinápticas) de los terminales sensoriales, pero están fundamentalmente regulados por un conjunto de interneuronas del ganglio, cuyos patrones de actividad producen varios comportamientos, como el repliegue branquial correlativo al tacto de la piel del sifón. Se observan cincuenta interneuronas, cada una de ellas haciendo contacto con un cuarto de lasrestantes interneuronas; se indican aproximadamente 300 interacciones sinápticas. Todo el círculo sensoriomotor está dentro del cuerpo de la *Aplysia* (arriba a la derecha) en constante acoplamiento con el entorno a través de su comportamiento autoinducido. *Adapatado de Zecevic et al. (1989), p. 3668.*

sea a través de una ruta anatómica distinta. Considérese por ejemplo el caso del sistema visual de los mamíferos. Podemos estar seguros del tan conocido flujo desde la retina a la así llamada "estación de relevo" (*relay station*) en el sistema visual, el tálamo dorsal (llamemos a esto región A) y luego del tálamo a la corteza visual primaria, (llamemos a esto B), y luego hacia otras regiones corticales. Lo que es menos discutido es que de acuerdo a la Ley de Reciprocidad, las conexiones desde B de vuelta a A, desde la corteza de vuelta *al* tálamo, son incluso más numerosas que aquellas de A a B (34, 39). Estas dinámicas bi-direccionales entre tálamo y corteza de las neuronas, no son tan sólo una sutileza anatómica. La *performance* visual de un animal se basa en este constante ir hacia atrás y hacia adelante, tal como lo muestra, por ejemplo, el fenómeno de la rivalidad binocular (36).

La dinámica neuronal que subyace a la tarea perceptuo-motora es entonces un asunto de redes, un sistema altamente cooperativo de dos vías, y no una secuencia informativa abstracta. Las densas interconexiones entre las subredes suponen que cada neurona activa operará como parte de un gran conjunto repartido en el cerebro, incluyendo tanto regiones locales como distantes. Por ejemplo, si bien las neuronas de la corteza visual tienen distintas respuestas ante "signos" específicos de estímulo visual (posición, dirección, contraste, etc.), estas respuestas sólo se dan en un animal anestesiado y con un alto grado de simplificación del entorno (interna y externamente). Cuando se han utilizado mejores condiciones sensoriales y se ha estudiado al animal mientras está despierto y activo, ha quedado cada vez más claro que las respuestas neuronales a los "signos" (estímulos) son altamente inestables y sensibles al contexto. Esto se ha demostrado, por ejemplo, en los efectos de la inclinación del cuerpo y la estimulación auditiva (37, 38, 39). Más aún, la respuesta característica de muchas neuronas en la corteza visual depende directamente de otras neuronas localizadas lejos de su campo receptivo (40): incluso un cambio de postura, a la vez que se mantiene idéntico el estímulo sensorial, altera las respuestas de las neuronas, lo que demuestra que incluso el supuesto flujo *motorium* está en resonancia con el *sensorium* (41).

Me permito seguir usando la visión como ejemplo, para retomar la discusión anterior y subirla un nivel en generalización, y hacer ver que en los últimos años la investigación de la visión ha llegado ser el estudio, no de una "reconstrucción" centralizada de una escena visual para beneficio de posteriores humanoides, sino aquella de un *patchwork* de modalidades visuales que incluye, al menos, forma, (forma, tamaño, rigidez), propiedades de la superficie (color , textura, reflejo especular, transparencia), relaciones espaciales tridimensionales (posiciones relativas, orientación tridimensional en el espacio, distancia) y movimientos tridimensionales (trayectoria, rotación). Ya es evidente que estos distintos aspectos de la visión son propiedades emergentes de subredes que compiten entre sí, que tienen un grado de independencia e incluso son anatómicamente separables, pero que están correlacionadas y trabajan juntas de modo que una percepción visual *es* esta coherencia (42).

Esta clase de estructura, para usar la metáfora de Minsky, es muy semejante a una "sociedad" de agentes, (43). Esta multiplicidad multidireccional es contraintuitiva, pero común a los sistemas complejos. Es contraintuitiva porque estamos acostumbrados al modo tradicional causal de ingreso-procesamiento-salida direccional. Nada sugiere, en la descripción anterior, que el cerebro funcione como una computadora digital, procesando información paso a paso; para el cerebro es muy difícil aplicar esta popular descripción de un sistema. En lugar de eso, a la red y a la arquitectura paralela le corresponden un tipo de operación distinta: hay un tiempo de "relajación" de señales que van y vienen, hasta que cada componente es puesto en actividad coherente. Por lo mismo, el ejercicio cooperativo toma cierto tiempo en completarse, y esto se hace evidente en que todo animal tiene su ritmo temporal natural. En el cerebro humano, este frenesí cooperativo toma alrededor de 200-500 milisegundos, el "ahora" (*nowness*) de una unidad perceptuo-motora (44, 45, 46). Al contrario de que lo que a primera vista pueda parecer —desde un punto de vista etológico o en nuestra propia introspección—, la vida cognitiva no es una corriente continua sino que se presenta

interrumpida por patrones conductuales que aparecen y desaparecen en el tiempo. Esta percepción de la neurociencia reciente —y de la ciencia cognitiva en general— es fundamental por cuanto nos libera de la tiranía de buscar una cualidad homuncular, centralizada, para el comportamiento de un agente cognitivo normal. Volveremos sobre esto más adelante.

Permítanme volver atrás un momento y reformular nuestra discusión sobre el ser cognitivo junto a aquello del ser molecular mínimo. Postulo que la neurociencia contemporánea —al igual que la célula biológica para el caso de la organización de la vida— da suficientes elementos para la concepción de la organización básica de un ser cognitivo en términos de *cierre operacional* (no interaccional) del sistema nervioso (2,17). Me refiero a "cierre" para destacar la cualidad autorreferencial de las redes de interneuronas y de las superficies percepto-motoras cuya correlación favorece. La calificación de "operacional" enfatiza que el cierre es usado en su sentido matemático de recursividad, y no en el sentido de encierro o aislamiento de la interacción, lo cual sería un absurdo, por supuesto. En términos más específicos, el sistema nervioso está organizado mediante operaciones de *cierre* de una red de subredes modular y recíprocamente relacionadas, dando origen a conjuntos de actividad coherente tales como:

i) patrones de correlación sensorimotora que son estables

ii) una conducta para todo el organismo como una unidad móvil en el espacio.

El *cierre* operacional del sistema nervioso produce un modo específico de coherencia, inscrustado en el organismo. Esta coherencia es un *ser cognitivo*: una unidad de percepción-movimiento en el espacio, de constantes senso-motoras medidas a través de la red de interneuronas. El paso a la cognición ocurre en el nivel de la entidad conductual, y no, como en el ser celular básico, como una entidad espacialmente limitada. La clave en este proceso cognitivo es el sistema nervioso través de su neuro-lógica. En otras palabras, el ser cognitivo es la manera en que el organismo, mediante

su actividad autoproducida, llega a ser una entidad distinta en el espacio, aunque siempre acoplada a su correspondiente medioambiente, del cual, no obstante, será siempre diferente. Un ser distinto coherente que, por el proceso mismo de constituirse (a sí mismo), configura un mundo externo de percepción y acción.

3.2 *Ser cognitivo y mundo perceptual*

La naturaleza de la *identidad* del ser cognitivo que acabamos de discutir, como la del ser celular básico, es de la *emergencia* a través de un proceso distributivo. Las propiedades emergente de un red de interneuronas son, sin embargo, enormemente más ricas y ameritan ser discutidas. Lo que deseo traer a colación aquí es la relativamente reciente (¡y aplastante!) conclusión de que muchos agentes simples con propiedades simples pueden ser reunidas, incluso al azar, para *dar nacimiento* a lo que aparece ante el observador como un todo múltiple e integrado, *sin* necesidad de una supervisión central. Ya hemos tocado este tema cuando discutimos la naturaleza del proceso autopoiético y el modelaje autómata celular, y más adelante cuando discutimos la constante emergencia y desaparición de los *ensembles* neuronales que está en el origen del comportamiento. Me gustaría tratar este asunto de manera más sustancial, dado que es crucial para toda mi argumentación. He basado mis conclusiones en estudios contemporáneos sobre una variedad de sistemas complejos inspirados en la biología. (2,47,48,49,50)

Permítanme volver a uno de los mejores ejemplos de esta conducta emergente: las colonias de insectos. Se han hecho muy populares en los estudios recientes sobre vida artificial. (AL) (51). Sus cualidades de "superorganismo" fueron entendidas como una metáfora por mucho tiempo, hasta que lentamente se retomaron en los años setenta (52) con algunos experimentos detallados que necesitaban un nivel de colonia global para sus explicaciones. Por ejemplo, algunos de los experimentos más ilustrativos envolucran una "sociotomía" seguida de las amplias propiedades reguladoras del conjunto. Después de separar una subcolonia, las nodrizas más

eficientes de la colonia *Neoponera apicalis* cambiaron radicalmente su status social, tendiendo más a la recolección que a la asistencia. Y ocurrió lo contrario en la colonia original, donde las nodrizas de más bajo nivel aumentaron su actividad asistencial. El experimento entregó evidencia de identidades configuradas con memoria, dado que las nodrizas volvieron a su estado anterior cuando la sociotomía fue interrumpida y los elementos separados regresaron a la colonia (53). Algunas conductas se pueden formalizar con herramientas que son bastante comunes para sistemas complejos, tales como ecuaciones de reacción-difusión para explicar la conducta recolectora de las hormigas (53). Efectos similares afectan la autodistinción espacial de esta colonia y su ontogenia.

Lo que impresiona especialmente en el caso de la colonia de insectos —al revés de lo que pasa en el cerebro— es que estamos prontos a admitir que (i) sus componentes separados son individuos y (ii) que no hay un centro o un ser localizado. Aun cuando el todo se comporta como una unidad y para el observador es como si hubiera un agente coordinador "virtualmente" presente en el centro. A esto me refiero cuando hablo de un sí mismo *no sustancial*: también se podría hablar de un ser *virtual*. Un patrón global coherente que emerge de simples componentes locales, que parece tener un localización central donde no se encuentra a nadie, pero que es esencial como nivel de interacción para la conducta de toda la unidad.

La adopción de estos modelos actuales, formalismos y estudios de casos de sistemas complejos (por ejemplo, propiedades emergentes a través de simples elementos coordinados) es, desde mi punto de vista, de gran importancia para la comprensión de las propiedades cognitivas. Introduce una alternativa explícita a la tradición computacional dominante en el estudio de las propiedades cognitivas, donde la idea central es el procesamiento de información externa sucesivamente elaborada para reconstituir una representación central (55,56). Este paradigma fundamental del programa de computador digital no sirve para la biología, ni para la IA.

Esto es evidente, tal como ya lo hemos discutido, para los neurobiológos que se ocupan del cerebro, pero también es cierto

para los investigadores de sistemas cognitivos artificiales, como lo han establecido claramente las escuelas *coneccionistas* actuales (49,57). Lo que encontramos en el cerebro es una promiscua recomposición de redes y sub-redes que no muestra evidencias de una descomposición estructurada de arriba abajo, cosa típica en el algoritmo computacional. De acuerdo con esto, uno de los primeros mensajes que se obtiene del estudio de redes de neuronas artificiales en modernos términos coneccionistas, es la ausencia de una distinción fundamental entre software y disco duro o, más precisamente, entre símbolos y no-símbolos. Es así como, todo lo que encontramos en las maquinarias de redes de neuronas artificiales son actividades relacionadas entre los conjuntos que marcan las regularidades y que llamamos sus conductas o *performances*. Podemos ver que algunos de estos conjuntos los describen de manera regular como programas, pero este es otro tema. Aun cuando estén artificialmente construidos, estos conjuntos emergentes no pueden ser llamados computacionales en el sentido que sus dinámicas no pueden ser formalmente especificadas como la implementación de un algoritmo de alto nivel. Las redes neuronales, incluso en sus detalles más sutiles, *no* son como una máquina de lenguaje, dado que simplemente no hay transición entre tales átomos operativos elementales y una semántica y el nivel emergente mayor en donde la conducta ocurre. Si la hubiera, la sabiduría clásica del computador respondería inmediatamente: Olvídese del disco duro, dado que no agrega nada significativo a la computación actual (excepto reducir el tiempo y el espacio). Por el contrario, en los modelos de red distributiva son precisamente estos "detalles" los que hacen posible un efecto global, por lo que marcan un profundo corte en la tradición de la IA (58,59). Por supuesto que esto refuerza una conclusión similar aplicable a las redes de neuronas naturales en el cerebro, tal como discutimos anteriormente.

He tratado este punto para advertir al lector sobre la fuerza acumulada en muchos años de dominación del idioma computacional, y su consecuente tendencia a identificar el ser cognitivo con algunos programas de computación o descripciones computacionales de alto nivel. Esto no sirve. El ser cognitivo *es* su propia

implementación: su historia y su acción son una sola pieza. Ahora, esto exige que aclaremos el segundo aspecto del ser que nos interesa tratar: su modo de relacionarse con el medioambiente.

La vida cotidiana es necesariamente una vida de agentes *situados*, que llevan adelante sus quehaceres enfrentados a una constante de actividades paralelas en su variados sistemas perceptuo-motores. La continua redefinición de los quehaceres no es de ninguna manera un plan, guardado en un repertorio de alternativas potenciales, sino una realidad dependiente de la contingencia, de la improvisación, y más flexible que planificada. Estar situado significa que la entidad cognitiva tiene, por definición, una perspectiva. Esto significa que no está "objetivamente" relacionada con su medioambiente, que es independiente del sistema de locación, títulos y apartados, actitudes e historia. En cambio, se relaciona con el medioambiente desde la perspectiva establecida por la constante emergencia de características del agente mismo y en términos del papel que juegan tales redefiniciones en la coherencia del sistema.

Un vez más, tal como hicimos con el ser celular mínimo, debemos hacer una diferencia tajante entre medioambiente y mundo. Y una vez más la manera de acoplamiento es doble. Por una parte, este cuerpo-en-el-espacio claramente sucede a través de las interacciones con el entorno del cual depende. Estas interacciones tienen la naturaleza de encuentros macrofísicos —transducciones sensoriales, fuerza del músculo y actuación, luz y radiaciones, etc.—, nada de qué sorprenderse. Sin embargo, este acoplamiento sólo es posible si los encuentros son asumidos *desde la perspectiva* del sistema mismo. Esto equivale, rigurosamente, a elaborar una *significación extra* relativa a esta perspectiva.

Todo lo que se encuentre deberá ser valorado de una u otra manera —gusto, disgusto, ignorancia— y recibido de una u otra manera —atracción, rechazo, neutralidad. Esta afirmación básica es inseparable de la manera en que el acto de acoplamiento se encuentra con una unidad perceptuo-motora en funcionamiento, y da lugar a una *intención* (incluso es tentador decir "un deseo"), esa cualidad única de la cognición viva (60).

Dicho en otros términos, la naturaleza del medioambiente adquiere para un ser cognitivo un curioso status: es la que *se presta* (*es lehnt sich an...*) a un excedente de significación. Como la improvisación de jazz, el medioambiente otorga la "excusa" para la "música" neural desde la perspectiva del sistema cognitivo involucrado. Al mismo tiempo, el organismo no puede vivir sin este acoplamiento constante y la igualmente constante emergencia de regularidades: de faltar la posibilidad de actividad asociaciativa, el sistema se convertiría en un simple fantasma.

Por ejemplo, la luz y el reflejo (entre otros parámetros macrofísicos como bordes y texturas, pero simplifiquemos para argumentar mejor), se prestan a una gran variedad de gama de colores, dependiendo del sistema nervioso involucrado en tal encuentro. Durante sus respectivos caminos evolutivos, los peces teleósteos, pájaros, mamíferos e insectos han generado diferentes gamas de colores, no solamente con distintas significaciones conductuales sino además con diferentes dimensiones de manera que no se trata de mayor o menor resolución de los colores (61). Es demostrable que el color no es una propiedad que se pueda "recuperar" de la "información" medioambiental de una manera precisa. El color es una dimensión que se muestra sólo en el diálogo filogenético entre un medioambiente y la historia de un ser autónomo activo, lo que en parte define lo que puede ser entendido como un medioambiente. La luz y el reflejo entregan un modo de asociación, una perturbación que gatilla, que da lugar a una enorme capacidad in-formativa de las redes neuronales para constituir correlaciones senso-motoras y, en consecuencia, poner en acción su capacidad para imaginar y presentarse. Es solamente *después* que ha sucedido todo esto, después que un modo de asociación se hace regular y repetitivo, como los colores en nuestro —y otros— mundo, que decimos en tanto observadores, por una cuestión de facilidad lingüística, que los colores corresponden o representan algún aspecto del mundo.

Un ejemplo dramático y reciente de este excedente de significación y de la deslumbrante actuación del cerebro como generador de narrativas neurales lo entrega la tecnología de la así llamada "realidad virtual" (62). Se unen un casco provisto de dos cámaras

sobre ambos ojos y un guante o traje con transductores de movimiento, no mediante el usual enlace a través de medioambiente —familiar a todos nosotros— sino a través de una computadora. De manera que cada movimiento de la mano o el cuerpo es correaltivo a imágenes a través de principios de correlación que pueden ser elegidos a voluntad. Por ejemplo, cada vez que mi mano, que aparece como una mano icono "virtual" en mi imagen, apunta hacia un lugar, la imagen que sigue corresponde a un vuelo hasta el lugar indicado. La percepción visual y el movimiento, por tanto, dan lugar a regularidades propias de esta nueva manera de asociación perceptuo-motora. Lo que me parece más significativo aquí es la *veracidad* del mundo que aparece: después de experimentar la nueva situación por un corto tiempo (15 minutos o algo así), *habitamos* un cuerpo al interior de este nuevo mundo y la sensación es de un verdadero vuelo a través de las paredes o una zambullida en universos fractales. Esto es así a pesar de la pobre calidad de las imágenes, la baja sensibilidad de los sensores y la limitada cantidad de interreaciones entre las superficies sensoriales y las imágenes, producto de las limitaciones del programa de la computadora. En su completud, el sistema nervioso es un sintetizador de regularidades tan magnífico, que cualquier material básico le es suficiente como medioambiente para producir un mundo extraordinario.

Esta misma estrategia de estar-en situación de un agente, que es progresivamente provisto de módulos cada vez más ricos de autoorganización interna, se está convirtiendo en un productivo programa de investigación, incluso en un campo tan pragmáticamente orientado como el de la inteligencia artificial (63, 64, 65). Cito —de manera algo extensa— a R. Brooks, uno de los más connotados exponentes de esta tendencia:

Yo ... abogo por una manera distinta de crear Inteligencia Artificial:

- Debemos incrementar las capacidades de los sistemas inteligentes en cada paso del proceso y por tanto asegurar, automáticamente, que las piezas y sus conexiones son válidas.

- En cada paso deberíamos construir sistemas inteligentes completos que pudiéramos dejar perderse en el mundo real con sensaciones reales y acciones reales. Menos que esto nos ofrece un candidato con el que sólo podremos engañarnos a nosotros mismos.

Hemos estado trabajando en esta dirección y hemos construido una serie de robots móviles autónomos. Hemos llegado a una conclusión inesperada (C) y tenemos una hipótesis bastante radical (H).

(C) Cuando examinamos niveles de inteligencia muy simples nos encontramos con que las representaciones explícitas y los modelos del mundo simplemente se nos atraviesan en el camino. Lo que parece más adecuado es usar el mundo como su propio modelo.

(H) La representación es la unidad de abstracción equivocada en la construcción de las partes más voluminosas de los sistemas inteligentes.

La representación ha sido el tema central en el trabajo sobre Inteligencia Artificial en los últimos 15 años, sólo porque ha entregado una conexión entre los módulos y los ensayos de conferencias que de otro modo quedarían aislados (65, p.1)

Cuando se alcanza la síntesis del comportamiento inteligente de esta manera, con estricto apego a la viabilidad senso-motora de un agente, la noción de que el mundo es una fuente de información para ser representada, simplemente desaparece. La autonomía del ser cognitivo se presenta con toda claridad. En la propuesta de Brook, la criatura mínima asume varias actividades al mismo tiempo, mediante una regla de cohabitación entre las mismas. Esto es semejante a un camino evolutivo, a lo largo del cual las subredes modulares se entretejen en el cerebro. Los resultados son aparatos dadores de sentido, verdaderamente inteligentes y autónomos, antes que débiles procesadores de información dependientes de un medioambiente preotorgado o de un plan perfecto.

Es interesante agregar que, en este artículo, Brooks también rastrea el origen de lo que describe como “la decepción de la IA” por la tendencia a la abstracción (no solamente en la IA, sino también en el resto de las ciencias cognitivas), por ejemplo, en la fabricación

de percepciones situadas y habilidades motoras. Tal como he sostenido aquí (y como discute Brooks por sus propias razones), tal abstracción no da cuenta de la esencia de la inteligencia cognitiva, *que reside sólo en su encarnación.* Es como si uno pudiera separar los problemas cognitivos en dos partes: una que puede ser resuelta mediante la abstracción y otra que no. La segunda incluye las capacidades motoras y de percepción-acción de agentes situados en medioambientes no especificados. Cuando nos acercamos al problema desde esta perspectiva autosituada, no hay un lugar donde la percepción pudiera entregar una representación del mundo a la manera tradicional. El mundo se muestra mediante la representación de las regularidades perceptuo-motoras. "Así como no hay representación central tampoco hay sistema central. Cada rango de actividad conecta directamente percepción y acción. Es el observador de la Creatura quien atribuye una representación central o un control igualmente central. La creatura en sí misma no tiene ninguno: es una colección de conductas competitivas. Del caos local producido por sus interacciones, surge, a los ojos del observador, un modelo de conducta coherente" (64. p. 11).

Para concluir, los dos puntos que he intentado aclarar en esta sección dedicada al ser cognitivo, son los siguientes: primero, he tratado de describir la naturaleza de su *identidad* como un cuerpo en movimiento-y-espacio, a través del cierre operacional de la red interneuronal. Esta actividad es observable como múltiples sub-redes, actuando en paralelo y entretejidas en complejos *bricolages*, generando de manera sucesiva e incesante modelos coherentes que se manifiestan como comportamientos. Segundo, he intentado aclarar cómo esta dinámica emergente, paralela y distributiva, es inseparable de la *constitución del mundo*, que no es otra cosa que el excedente de significado y de intenciones que aparece en los comportamientos situados. Si las conexiones con el ambiente físico son inevitables, lo que hace único al ser cognitivo es esta génesis constante de significado. O, invirtiendo la descripción, lo que hace único al ser cognitivo es su *falta* constitutiva de significación, que debe ser resuelta en el enfrentamiento permanente

con las perturbaciones y rupturas propias de la vida perceptuo-motora. La cognición es acción referida a *aquello que falta*, visto desde la perspectiva de un ser cognitivo que siente aquella falta.

4. Organismo: un trenzado de seres virtuales

4.1 *El sí mismo como persona*

Incluso si nos gustaran estas ideas sobre el sí mismo no sustancial, tanto a nivel celular básico como al nivel más elaborado de la cognición, "deseamos eximirnos a nosotros mismos (deseamos exirmirnos a nosotros *mismos*). El problema es que *parece* que al menos *somos* claramente diferentes: erguidos, centrados, dirigidos de manera global" (66, p.6). Es por esto que nos sentimos forzados a proyectar un centro centralizado o agente, una entidad homuncular dotado de un alma o una vaga idea de "el ser como proceso".

Creo que la gran novedad de nuestro recién adquirido y todavía fragmentario, entendimiento de las propiedades emergentes en los procesos de redes distributivas, se encuentra precisamente en que son metáforas *fuertes*, ejemplares, de qué es un sí mismo no sustancial: un vacío coherente que no es encontrable en ninguna parte, y, *sin embargo,* es capaz de otorgar condiciones para una asociación. Subrayo lo fuerte de la metáfora porque, de faltar estos ejemplos, esta aparente paradoja de una no-localización responsable de la designación como totalidad, se convierte en una contradicción. A menos que esta aparente paradoja sea incluida en este metanivel constructivo, rápidamente recaemos en la discusión tradicional acerca de la existencia versus la no existencia del ser, la persona, holismo y cosas parecidas. La novedad que aquí se ofrece es que cambiamos de nivel a través de un pasaje de dos vías: "hacia arriba" como propiedades emergentes de los elementos constitutivos, y "hacia abajo" como restricciones a las interacciones locales debidas a la coherencia global. Por lo tanto, un ser no substancial puede actuar como si estuviera presente, como una conexión virtual.

A medida que vamos reconociendo la naturaleza no sustancial/vacía del sí mismo en distintas "regiones", más

sospechamos de nuestro sentimiento del "yo" como verdadero centro. O somos únicos, en oposición a todo lo demás encontrable en el mundo viviente y natural, o nuestra sensación inmediata de un ser personal y central es el mismo tipo de ilusión de un centro que encontramos en los análisis que hemos hecho de los seres cognitivos celulares y básicos, en sus respectivos niveles.

No es necesario agregar que prefiero claramente la segunda alternativa. Lo que podemos llamar "yo", nosotros mismos, puede ser analizado como algo que surge de la capacidad lingüística del ser humano y su capacidad particular para autodescribirse y narrar. Como enseñan las evidencias neurofisiológicas, el uso del lenguaje es otra capacidad modular en co-habitación con todo lo demás que podemos conocer. Nuestro sentido de un "yo" personal puede ser entendido como una permanente narración interpretativa de algunos aspectos de las actividades paralelas en nuestra vida cotidiana, por ende, de los constantes cambios en las maneras de atender propias de nuestro ser (nosotros mismos) central. Y de esto deriva la relativa fragilidad de tales construcciones narrativas (67).

Dado que este "yo" narrativo se constituye necesariamente a través del lenguaje, entonces este sí mismo personal está unido a la vida *social*, porque el lenguaje no puede operar sino como fenómeno social. En el hecho, uno podría ir un poco más lejos: probablemente el "yo" no sustancial/vacío sea el puente entre el "cuerpo corpóreo" (corporeal body) común a todos los seres con sistema nervioso y la dinámica social en la que se desenvuelve el ser humano. "Yo mismo" no es solamente privado ni solamente público, pero es ambas cosas a la vez. Y lo mismo sucede con los tipos de narrativas que acompañan a los "yoes", tales como valores, hábitos y preferencias. Desde un punto de vista lógico funcionalista, el "yo" puede ser entendido *para* la interacción con otros, *para* la creación de la vida social. De estas articulaciones surgen las propiedades emergentes de la vida social, donde los "yoes" insustanciales/vacíos son los componentes básicos. Por tanto, siempre que encontramos regularidades como las leyes o los papeles

sociales y los concebimos como algo dado desde afuera, caemos en la misma falacia de otorgar a cualquier propiedad emergente una identidad substancial, en lugar de verlas como propiedades emergentes de un complejo proceso distributivo, mediado por las interacciones humanas. Estas emergencias sociales pueden ser proyectadas como puntos de referencias "exógenos", tal como se hace en las sociedades tradicionales, pero pueden ser igualmente deconstruidas mediante el tipo de argumento que hemos utilizado aquí (68).

Lo peculiar de este "ser personal" (personal self) —y que define bien lo que es el nivel mental o psicológico— es que a través del lenguaje hay una relativa autonomización del sí mismo cognitivo básico (*basic cognitive self*), tal como la percepción-acción entrega un completo y nuevo campo de significación autónoma, relativas a las interacciones macrofísicas ambientales. Esta narrativa, en sí misma, se convierte en un mundo *para* un sujeto en su sentido más tradicional y literal, la completa autonomización del registro imaginario.

Es interesante, incluso si aceptamos una reinterpretación del "yo mismo" como virtual —a través del lenguaje y las propiedades distributivas emergentes— observar que nuestra inclinación natural en la vida cotidiana es seguir adelante como si nada hubiera cambiado. Esta es la mejor evidencia de que el proceso de autoconstitución está tan arraigado que para deconstruirlo no es sólo cuestión de contar con un análisis convincente. La fuerza activa orientada a la constitución de identidad es profundamente inalterable. Explorarla y deconstruirla es esencialmente un asunto de aprendizaje y transformación sostenida. Esta es una de las razones por las que la tradición filosófica continental, especialmente representada por M. Merleau-Ponty (69), ha insistido en que nuestra condición debe ser puesta en una doble perspectiva: por una parte necesitamos entender nuestras condiciones como procesos corporales: por otra, somos también una existencia que *está ahí*, un *Dasein*, constituido como una identidad, que no puede simplemente salir y mirar desde una perspectiva incorpórea para entender cómo llegó ahí donde está. Para Heidegger, la condición humana no es la de un organismo

animal con "algo" agregado a él. La animalidad es más bien un *Dasein* empobrecido: el animal es "pobre en su mundanería" (Weltarm) (70). En otro lugar me he referido extensamente a esta doble perspectiva de aproximación a la experiencia humana y los mecanismos cognitivos (71).

Es evidente que, tratar en profundidad estos temas nos llevaría demasiado lejos. Como en el caso de la discriminación inmunológica entre ser y no ser, estoy aquí entregando lo que obligadamente será un frustrante esbozo abierto a toda clase de objeciones teóricas y empíricas. No obstante, debo decir que mi propósito, al tratar el sí mismo como "yo", es enfatizar *la continuidad* del motivo (*motif*) que discutimos ampliamente en el caso de los seres cognitivos celulares y básicos. Como un fractal, este motivo se repite una y otra vez en los variados sí mismos regionales del organismo, incluso cuando la articulación entre ellos, hasta ahora, está muy lejos de ser satisfactoria. Pero desde mi punto de vista es muy importante reconocer cómo este motivo aparece tanto en nuestra experiencia más íntima e inmediata como en los niveles más profundo de la vida y el cuerpo. Sólo de esta manera podemos evitar la dispersión de los sí mismos de un organismo en categorías dispares y, por lo tanto, evitar la dispersión de lo que es una totalidad que incluye desde las células hasta las ideologías, en campos pulverizados por la separación.

4.2 *La dialéctica doble de los organismos*

El organismo como tal, entonces, no puede ser abordado como un proceso único. Nos vemos forzados a descubrir "regiones" que se entretejen de complejas maneras y, en el caso de los humanos, se extiende más allá de los estrictos límites del cuerpo, hasta el registro socio-lingüístico.

Además, lo que he argumentado es que detrás de esta malla de variados sí mismos que llevamos con nosotros, todos estos sí mismos comparten una lógica común y fundamental, aun cuando difiera en su especificidad. Es el caso de lo que Wittgenstein hubiera

llamado "parecidos familiares": antes que cualquier característica común a todas las instancias, nos encontramos con un *racimo* de características superpuestas. También podemos hablar de este racimo de características comunes como una *dialéctica* compartida, ya que nos encontramos ante procesos de doble cara, donde la co-definición es un punto central. En rigor, propongo que la dialéctica orgánica del ser es un asunto de dos niveles: tenemos por un lado la dialéctica de la identidad del ser; del otro la dialéctica a través de la cual esta identidad, una vez establecida, produce un mundo desde un medioambiente. La identidad y el conocimiento se relacionan mutuamente como dos caras de un proceso único que forma el núcleo de la dialéctica de todos los sí mismos.

Primero, una *dialéctica de identidad* establece un agente autónomo, un *para* sí (*pour soi*). Esta identidad se establece mediante el ensamble de dos términos:

(i) un término *dinámico* referido a una conjunción de componentes interactuando en red, capaz de desarrollar propiedades emergentes: mallas metabólicas, conjunciones neuronales, redes de anticuerpos, recursividad lingüística;

(ii) un término global referido a las propiedades emergentes, una totalidad que condiciona (en sentido descendente) los componentes de la red: membranas celulares, cuerpo senso-motor en el espacio, discriminación ser/no ser, "yo" personal.

Estos dos términos están realmente en una relación de co-definición. Por una parte el nivel global no puede existir sin el nivel de red, dado que se produce a través de este; por otra, el nivel dinámico no puede existir ni operar como tal si no es contenido y alojado en una unidad adecuada que lo haga posible.

Segundo, una *dialéctica del conocimiento* establece un mundo de significación cognitiva *para* esta identidad. Esto sólo puede surgir desde la perspectiva que la misma identidad entrega, lo que agrega un excedente de significación a la interacción del medioambiente propio de las partes constitutivas.

El punto central, entonces, es que el organismo produce y especifica su propio campo de problemas y acciones por ser "resueltos"; este campo cognitivo no existe "allá afuera", en un medioambiente que actuaría como pista de aterrizaje de organismos que de una u otra manera irían cayendo sobre el mundo. En lugar de eso, los seres vivientes y sus mundos están en relación unos con otros mediante *especificaciones mutuas* o *co-determinaciones*. Por tanto lo que describimos como regularidades medioambientalmente significativas no son características que han sido internalizadas, como asumen la tradición representacional dominante en las ciencias cognitivas y el adaptacionismo en la biología evolucionista. Las regularidades medioambientales son el resultado de una historia conjunta, una congruencia que se desprende de una larga historia de co-determinación. En palabras de Lewontin (73), el organismo es tanto el sujeto como el objeto de la evolución.

Este segundo nivel de la dialéctica del organismo, entonces, se establece también mediante la conjunción de dos términos:

(i) un término de *significación* referido a la necesaria emergencia de un excedente de significación (meaning), propio de la perspectiva del ser constituido: semántica celular, percepción y acción conductual, ser/no ser como asertos somáticos, identidad personal,

(ii) un término de *enlace* referido al permanente y necesario empotramiento y dependencia del ser respecto de su medioambiente, ya que sólo mediante dicho enlace puede crear su mundo: leyes físico-químicas para el mundo celular, propiedades físicas microscópicas para el comportamiento cognitivo, interacción molecular para el ser inmune, intercambios socio-lingüísticos para nuestro ser subjetivo.

Doble dialéctica: la naturaleza de una identidad y la naturaleza de una relación con el mundo. Doble paradoja: autoproducción por contenimiento dependiente; autonomía de conocimiento a través del enganche con el medioambiente. Ambas dialécticas dan lugar a la naturaleza cambiante del organismo, siempre formándose

a sí mismo e informando dónde está, y siempre implicado en los antecedentes de donde proviene. Organismos, fascinantes redes de sí mismos no sustanciales/vacíos, nada más ni nada menos que existencias circulares, de niveles múltiples y final abierto, siempre movidos por la falta de significación que ellos mismos engrendan al afirmar su presencia.

REFERENCIAS BIBLIOGRAFICAS

1. Lewontin, R. (1982), *The Dialectical Biologist*, MIT Press, Cambridge.
2. Langton. C. (Ed.) (1989), Artificial *Life*, Addison-Wesley: Redwood City.
3. Maturana, H. and F. Varela (1973), *De máquinas y seres vivos. Uno teor!a de la organización biológica,* Editorial Universitaria: Santiago de Chile.
4. Maturana, H. and F. Varela (1980), *Autopoiesis and cognition: The realization of the living, D. Reidel: Boston. [Boston Studies in the Philosophy of Science, vol. 42].*
5. Varela, F., H. Maturana, and R. Uribe (1974), Autopoiesis: the organization of living system, its characterization and a model, *Biosystems 5:* 187-195.
6. Fleischaker, G. (1988), *Autopoiesis: System logic and the origin of life, Ph.D. D*issertation, Boston University, Boston, MA.
7. Margulis. L. (1981), *Symbiosis in Cell Evolution.* W.H. Freeman, San Francisco.
8. Margolis L. and D. Sagan (1986), Origins of *Sex.* Yale Univ. Press, New Haven.
9. Gardner, M. (1971), On cellular automata, self-reproduction, the Garden of Eden, and the game «life», *Sci. Amer.* 224:112.
10. Wolfram, S. (1986), *Theory and Applications of Cellular Automata,* World Scientific, Singapore.
11. Toffoli, T. (1987), C*ellular Automata Machines.* MIT Press, Cambridge.
12. Deamer, D. and G. Barchfeld (1982), Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions, *J. Molec Evol.* 18: 203-206.

13. Lazcano, A. (1986), Prebiotic evolution and the origin of cells, *Treballs Society Catal. Biol.* 39: 73-103.
14. Baeza, I. M. Ibañez, A. Lazcano, C. Santiago, C. Arguello, C. Wong, and J. Oró, Liposomes with polyribonucleotidles as models of precellular systems, *Origins of Life* 17:187-199.
15. Deamer, D.W. (1985), Role of amphillic compounds in the evolution of mebrane structure on the early Earth, *Origins of Life* 17: 3-25.
16. Luisi, L. and F. Varela (1989), Self replicating micelles: A minimal version of a chemical autopoietic system, *Origins of Life 19:* 633-643.
17. Varela, F. (1979), *Principles of Biological Autonomy.* North-Holland/ Elsevier, New York.
18. Varela, F. (1988), Structural coupling of simple cellular automauta: On the origin of meaning. *In: E.* Secarz, F. Celada, N.A. Mitchinson, and T. Tada, *The Semiotics of Cellular Communication in the Inmune System, NATO ASI* Series, Vol. 1123, Springer-Verlag, New York, pp.151-161.
19. Castoriadis, C. (1987), L'état du sujet aujourd'hui, *Topique* 38: 7-39.
20. Margulis, L. and K. Schwartz (1988), *Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth,* W.H. Freeman, New York.
21. Buss, L. (1987). *The Evolution of Individuality,* Princeton Univ. Press, Princeton.
22. Banner, J.T. (1988), *The Evolutio of Complexity, Princ*eton Univ. Press, Princeton.
23. Varela, F., B. Dupire, and A. Coutinho, (1988), Cognitive networks: Immune, neural and otherwise. I*n:* A. Perelson (Ed.), *Theoretical Inmunology, Vol.* 2. (SF1 Series on Complexity), Addison-Wesley, New Jersey, pp. 359-375.
24. Vaz, N. and F. Varela (1978), Self and nonsense: An organism-centered approach to immunology, *Medical Hypothesis 4:* 231-267.
25. Coutinho, A., L. Forni, 0. Holmberg, F. Ivars, and N. Vaz (1984), From an antigencentered, clonal perspective on immune responses to an organism-centered network perspective of autonomous activity in a self-referential immune system, *Immunol. Revs. 79:*151-168.
26. Lundqvist, I., A. Coutinho, F. Varela, and 0. Holmberg (1989), Evidence for the functional dynamics in an antibody network, Proc. *Natl. Acad. Sci. (USA)* 86*:* 5074-5078.
27. Varela, F. and A. Coutinho (1991), Second generation immune networks. *Immunol. Today. I*n press.
28. Coutinho, A. (1989), Beyond clonal selection and network, Immunol. *Revs.* 110: 63-87.

29. Varela, F., A. Andersson, G. Dietrich, A. Sundblad, 0. Holmberg, M. Kazatchkine and A. Coutinho, The population dynamics of natural antibodies in normal and autoimmune individuals, Proc. Na*tl. Acad. Sci. (U.S.A.),* In Press.
30. Varela, F. and S. Frenk (1987), The organ of form: Towards a biological theory of shape, *J. Soc. Biol. Struct.* 10: 73-83.
31. Carew, T. and C. Sahley (1983), Invertebrate learning and memory: from behavior to molecules, *Ann. Rev. Neurosci. 9:* 435-487.
32. Zecevic, 0., J. Wu, L. Cohen, J. London, H. Höpp, C. Falk (1989), Hundreds of neurons in the Aplysia abdominal ganglion are active during the gill-withdrawal reflex, *J.* Neurosci. 9:3681-3689.
33. John, E.R., Y. Tang, A. Brill, A.B. Young, and K. Ono (1986), Double-labeled metabolic maps of memory, Sci*ence* 233:1167-1175.
34. Singer, W. (1977), Control of thalamic transmission by corticofugal and ascending reticular pathways in the visual system, *Physiol. Rev.* 57: 386-420.
35. Steriade, M. and M. Deschenes (1985), The thalamus as a neuronal oscillator, *Brain Res. Rev.* 8: 1-63.
36. Varela, F. and W. Singer (1987), Neuronal dynamics in the visual cortico-thalamic pathway revealed through binocular rivalry, Ex*p. Brain Res.* 66: 10-20.
37. Horn, G. and R. Hill (1968), Modifications of receptive fields of cells in the visual cortex occurring spontaneously and associated with bodily tilt, N*ature* 221: 186-188.
38. Fishman, M.C. and P. Michael (1973), Integration of auditory information in the cat's visual cortex, Vision Res*earch* 13:1415-1419.
39. Morell, F. (1972), Visual system's view of acoustic space, Nature 238: 44-46.
40. Allman, J., F. Meizen, and E. McGuiness (1985), Non-classical receptive field properties, *Ann. Rev. Neuroscien. 8:* 407-430.
41. Abeles, M. (1984), Local Cortical Circuits, Springer Verlag, Berlin.
42. DeYoe, E. and D.C. Van Essen (1988), Concurrent processing streams in monkey visual cortex, *Trends Neurosci.* 11: 219-226.
43. Minsky, M. (1987), *The Society* of Mind, Simon and Schuster, New York.
44. Llinás, R. (1988), The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function, *Science* 242: 1654-1664.
45. Gevins, A., R. Schaffer, J. Doyle, B. Cutillo, R. Tannehill, and S. Bressler (1983), Shadows of thought: shifting lateralization of

human brain electrical patterns during a brief visuomotor task, *Science* 220: 97-99.
46. Varela, F., A. Toro, E. John, and E. Schwartz (1981), Perceptual framing and cortical alpha rhythms, *Neuropsychologia*, 19: 675-686.
47. Goodwin, B. and P. Saunders (Eds.) (1989) Theoretical Biology: *Epigenetic and Evolutionary Order* from Complex *Systems,* Edinburgh University Press, Edinburgh.
48. Farmer, J., A. Lapedes, N. Packard, and B. Wendroff (Eds.) (1986), Evolution, *Games and Learning*, North-Holland, Amsterdam.
49. McClelland, J. and D. Rummelhart (1986), *Parallel Distributed Processing: Studies on the Microstructure of Cognition*, 3 vols., MIT Press, Cambridge.
50. Wolfram, S. (1984), Cellular automata as models for complexity, Nature 311: 419-424.
51. Pasteels, J. and J. Deneubourg (1987), From Idividual to Collective Behavior in Social Insects, Birkhäuser, Basel.
52. Wilson, E.O. (1971), The Insect Societies, Harvard Univ. Press, Cambridge.
53. Fresnau, D. and J. Lachaud (1985), La régulation sociale: données préliminaires sur les facteurs individuels contrôlant l'organisation des tâches chez Neopo*nera apicalis, Actes Coll. Insects Sociaux* 2:185—I 93.
54. Deneubourg, J., S. Aron, S. Goss, J. Pasteels, and C. Duerinck (1986), Random behavior, amplification processes and number of participants: how they contribute to the foraging properties of ants. In: Farmer *et als.* (Eds.), *Evolution, Games and Learning, op. cit.*
55. Newell, A. (1980), Physical symbol systems, Cognitive *Scien.* 4:135-183.
56. Plyshyn, Z. (1984), Computation ad Cognition: Towards a Foundation foe Cognitive Science, MIT Press, Cambridge.
57. Grossberg, S. (1984), *Studies of Mind and Brain*, D. Reidel, Boston. [Boston Studies in the Philosophy of Science, vol. 70].
58. Smolensky, P. (1988), On the proper treatment of connectionism, *Beh. Brain Sci.* 11: 1-74.
59. Dennett, D. (I990a), Mother nature versus the walking encyclopedia: A westen drama, in: Ramsey, S., D. Rummelhart, and S. Stich (Eds.), *Philosophy and Connectionist* Theory, (forthcoming).
60. Dennett, D. (1987), *The Intentional Stance*, MIT Press, Cambridge.
61. Thompson, E., A. Palacios, and F. Varela (1991), Ways of coloring:

Comparative color vision as a case study in cognitive science, *Beh. Brain Sci.* In press.
62. Fisher, S. (1990), *In*: C. Hattinger (Ed.), Virtuelle *Welten,* Linz.
63. Agree, Ph. (1988), *The Dynamic Structures of Everyday Life*, Report No. AI-TR 1085, MIT Artificial Intelligence Lab., Cambridge.
64, Brooks, R.A. (1986), Achieving artificial intelligence through building robots, *A.I. Memo 899*, MIT Artificial Intelligence Laboratory, May 1986.
65, Brooks, R.A. (I 987), Intelligence without representation, MIT Artificial Intelligence Report, Cambridge, MA.
66. Dennett, D. (l990b), Review of Ch. Langton (Ed.), Artificial Lif*e, Biology Phylos.* lo press.
67. Humphreys, N. and D. Denett (1989), Speaking for ourselves: An assessment of multiple personality disorder, Raritan 9: 68-98.
68. Dupuy, J.-P. and F. Varela (1990), Understandings of Origins. *In:* Varela, F. and J.P. Dupuy (Eds.), *Understanding Origins: Contemporary Ideas on the genesis of life, mind and society*, Kluwer, Boston, In press.
69. Merleau-Ponty, M. (1952), Phénoméologie de la Perception, Gallimard, Paris.
70. Heidegger, M. (1983), Die Grundbegriffe der *Metaphysik, Gesamtausgabe* t.29/30, Klostermann, Frankfurt.
71. Varela, F., F. Thompson, and F. Roach (1991), *The Embodied Mind:* Cognitive Science and Human Experience, MIT Press, Cambridge.
72. Wittgenstein, L. (1972), Philosophical Investigations, Basil Blackwell, Oxford.
73. Lewontin, R. (1983), The organism as the subject and object of evolution, *Scientia* 118: 63-82.

EL CUERPO EVOCADOR
Una relectura de la inmunidad

Si la identidad sin centro es un concepto clave en este libro, una de sus manifestaciones más notables es la identidad somática del cuerpo, donde el sistema inmunitario juega el rol estelar. Esta perspectiva ha quedado al margen de una reflexión más amplia en gran parte debido a la persistencia de un pensamiento sobre la inmunología moderna centrado en los mecanismos de defensa. Este texto propone una relectura desde el interior de la inmunidad, y sobre todo en sus repercusiones sobre las múltiples asonancias de la distinción yo/no-yo, que tanto ha interesado a la tradición psicoanalítica: no es casualidad que los editores de Nouvelle Revue de Psychanalyse *fueran quienes me solicitaron este texto.*

En un sentido muy primordial, la intimidad más íntima es la de nuestra propia carne, el habitáculo de nuestro cuerpo en su identidad encarnada. Allí se encuentra el *locus* dinámico (extraño por ser íntimo) donde puede producirse un yo corporal.

Semejante postura implica y exige reapreciar el cuerpo humano: tratar de considerar al organismo como portador de un primer registro del Ser. Esta dimensión "óntica" del cuerpo ha sido, a nuestro entender, subestimada de manera evidente, en beneficio de las dimensiones socio-histórica y lingüística de la experiencia. A menudo, el psicoanálisis, en particular, ha considerado que esta primera y primordial dimensión era reducible a la segunda; en todo caso, ha actuado como si asumiera esta posición.

(Varela,F. and A.Cohen (1989), Le corps évocateur: une relecture de l'immunité, *Nouvelle Revue de Psychanalyse* N°40, Automne 1989, pp.193-213.)

Este trabajo se inscribe en el marco de un proyecto esencial, pero que sólo está en vías de realización: una "evocación corporal del Ser (*body's recollection of Being*)[1]". Se trata, al interior de y gracias al conjunto de nuestra experiencia, de continuar la crítica de la tradición racionalista en la cual el tema del conocimiento se estructura en torno al "ideal de la sutura", y todo conocimiento es "espejo de la naturaleza"[2].

El enlace entre el Ser y la existencia es nuestro cuerpo. Este enlace, lo más íntimo de nosotros, es un legado primordial, un orden implícito que informa toda experiencia y todas aquellas actividades más familiares y mundanas que calificamos como humanas. La visión que se nos da es la de un conjunto de procesos biológicos encarnados.

Esta condición corporal no es, por cierto, ni ahistórica, ni atemporal, ni está determinada por un programa genético, en cierta manera fijo, como lo cree cierta racionalidad científica. Es una condición que sólo puede manifestarse al interior del lenguaje y de la historia, pero que no está determinada por lo histórico *per se*. Es en los límites de nuestra escritura de nosotros mismos que nos acercamos a la trascendencia.

Para abordar esta comprensión previa, es imperativo tomar en cuenta los *dos* polos de su condición: por una parte, su naturaleza biológica y sus raíces en la vida animal, por otra, el contexto histórico-lingüístico en que uno se encuentra ubicado aquí y ahora, que es revelado necesariamente por una actividad hermeneútica cualquiera. Estos dos polos "se habilitan" mutuamente: uno se remite

---

1 En este sentido *Fragments for a History of the Body*, Zone, Nueva York, 3 vols,1988, es una excelente fuente reciente de trabajo y contiene una bibliografía anotada. Tomamos la frase citada del libro de D.M. Levin *The Body's Recollection of Being*, Routledge, Nueva York, 1985; ver también *The opening of Vision*, Routledge, Nueva York, 1988, de donde se debe mucha de la inspiración de esta parte del trabajo. También aprovechamos la lectura de M. Berman, *Coming to our Senses: The body in the hidden history of the West*, Simon and Schuster, Nueva York, 1989 y de C. Castoriadis, "L'état du sujet aujourd'hui ", *Topique*, nº38, 1986, pp 7-39.

2 J. Dor, *L'A-scientificité de la Psychanalyse*, vol. 2, Éditions Universitaires, París, 1988, pp 75-77. R. Rorty, *Philosophy as the Mirror of Nature*, Princeton UP, 1981.

a los límites de lo histórico, el otro pone en evidencia la forma en que se constituye el discurso biológico. Ambos se entrecruzan en la vivencia del cuerpo, en un cuerpo que es nuestra morada más íntima, y del que toda acción extrae su posibilidad. Esta condición bipolar del cuerpo primordial la designamos aquí a través de la expresión *el* entrelazado del cuerpo vivido[3], tomada de los trabajos visionarios sobre la evocación del cuerpo, de Merleau-Ponty.

Para que este enlace de nuestro cuerpo primordial se manifieste, es necesario estar dispuesto a escucharlo. Una vez más, esta escucha se activa a partir de dos polos. En primer lugar, mediante la deconstrucción de la representación dominante del cuerpo como mecanismo exterior, puramente natural, separado y distinto de nuestra experiencia. Esta es una tarea de revisión al interior de las ciencias naturales que son responsables de esta concepción. En segundo lugar, mediante el cultivo continuo de prácticas que pueden revelar esta comprensión previa al interior de nuestra experiencia, haciendo posible así el enlace progresivo con nuestro cuerpo primordial. En este sentido, toda "práctica de transformación del sujeto"[4] se equilibra simétricamente con la tarea de la revisión de las raíces biológicas del cuerpo y del conocimiento.

Ahora podemos precisar la intención principal de este texto: presentar un ejemplo de una revisión de una de las identidades fundamentales de nuestro cuerpo, la inmunidad. Insistimos en el hecho de que sólo se trata de un ejemplo. Se podrían haber elegido otras dimensiones, como la organización celular, la forma multicelular o el sistema neurocognitivo. Todas estas dimensiones se entrecruzan y coexisten y nos proponen una revisión de la biología

---

3 M. Merleau-Ponty, Le visible et l'invisible, Gallimard, 1964, p. 172

4 Se vuelve a tomar la conocida fórmula de Foucault, en *Le Souci de Soi*, Gallimard, París, 1987. Es verdad que el psicoanálisis es una de esas prácticas, pero se piensa en otras. Una práctica diferente e importante es la técnica de la atención, desarrollada en la tradición budista. Ver por ejemplo, Y. Yasua, *The Body*, SUNY, New York, 1988, y F. Varela y E. Thompson, *The Embodied Mind: Cognitive science and human experience*, MIT Press, 1991.

moderna que, como se verá a continuación, nos llevará a trazar las raíces del sujeto dividido hasta las identidades de su cuerpo[5].

1. Deconstrucción de la inmunidad como defensa

Uno de los fundamentos de la identidad del cuerpo, que se encuentra en el origen mismo de nuestra identidad molecular y celular, es el conjunto de mecanismos que se denomina *inmunidad* y el establecimiento de la discriminación del yo y de lo que no pertenece al yo[6]. Aquí proponemos otra lectura a través del enfoque utilizado para delimitar la identidad inmunitaria que puede extenderse a los fenómenos neurocognitivos y a otras identidades del cuerpo. Sin embargo, como este último tema es demasiado complejo para ser tratado en los límites de este texto, nos contentaremos con proporcionar al lector las ideas básicas que determinan esta concepción específica de una organización biológica que conduce *al surgimiento de una identidad autónoma.*

La función que se le atribuye a la inmunidad tiene que ver con la defensa de "sí mismo" contra las agresiones de elementos extraños. El problema de la identidad, de las fronteras y de lo que está en el interior ("lo íntimo", metafóricamente hablando), se formula aquí con toda su fuerza. Lo que esperamos del sistema inmunitario es que produzca defensas contra los invasores y células de vigilancia que puedan destruir a los agentes patógenos y preservar el yo de los invasores o de lo que no pertenece al yo. Todo texto de inmunología se dedica a definir la disciplina como el estudio del conjunto de estas "respuestas" inmunitarias.

---

5 La orientación general de esta revisión de la biología se encuentra en H. Maturana y F. Varela, *Autopoieis and Cognition,* Reidel, Boston, 1980; F. Varela, *Principles of Biological Autonomy*, Elsevier, New York, 1979; y F. Varela *Autonomie et Connaissance: Essai sur le vivant*, Seuil, 1988. Para una introducción reciente y accesible a este enfoque, ver H. Maturana y F. Varela, *The Tree of Knowledge*, New Science Library, Boston, 1987.

6 Esta parte se basa parcialmente en un texto redactado por Mark Anspach, que será retomado más adelante.

El discurso inmunológico gira en torno a metáforas *militares*, así como, en un momento dado, las ciencias que estudian el cerebro giraban en torno a metáforas informáticas. Nuestro trabajo inicial en la revisión de la inmunidad consiste en introducir un marco conceptual y metafórico sustancialmente diferente. Este marco subraya las capacidades *cognitivas* y de autoorganización de los procesos inmunitarios. (Aunque el término cognitivo pueda parecer demasiado fuerte, parece, no obstante, útil introducirlo, sólo para subrayar el contraste con la concepción militar de la inmunidad[7]).

Se podría comparar el cuerpo con la tierra, un medio ambiente compuesto por poblaciones de individuos variadas y altamente interactivas. En el caso que nos preocupa, los individuos son representados a través de los linfocitos que constituyen el sistema inmunitario, es decir, conjunto de células móviles que se diferencian entre sí mediante marcadores moleculares específicos, o anticuerpos "pegados" a la superficie de su membrana. Al igual que las especies vivientes de la biosfera, estas poblaciones de linfocitos estimulan o inhiben recíprocamente su crecimiento. Al igual que las especies de un ecosistema, éstos son de una diversidad asombrosa: los anticuerpos y las otras moléculas producidas por los linfocitos son, claramente, el conjunto más variado de moléculas fabricadas por el organismo (un millón de veces más que cualquier otro); mecanismos delicados aseguran la renovación constante y la diversidad de las células existentes.

Esta red de linfocitos se mantiene en armonía con su ecología natural, el ambiente *somático* del cuerpo, que determina las especies linfocitarias que estarán presentes. Pero, al igual que en la biosfera, los linfocitos existentes afectan de una manera radical cada perfil molecular del organismo. Es así como se elabora en nosotros, adultos, un modo nodal según el cual nuestra "manera de ser corporal" se organiza, una forma de identidad, que no es otra cosa que una

---

[7] Esta revisión de la inmunología ha sido publicada *in extenso* en otros trabajos, ver: F. Varela, A. Countinho, B. Dupire y N. Vaz (1988), "Cognitive networks: Immune, neural and otherwise" en, A. Perelson (Ed.) *Theoretical immunology*, vol 2. (CFI Series of Complexity), Alison Wesley, Nueva Jersey.

asociación cuerpo/inmunidad, formada, elaborada a lo largo de la vida como una configuración única.

Para cumplir un papel de defensa, el sistema inmunitario debe poseer propiedades típicamente cognitivas. En primer lugar, debe poseer la capacidad de *reconocer* diversos perfiles moleculares: los perfiles moleculares extraños (o antígenos, en jerga inmunológica), su naturaleza "extraña" y su potencial patógeno que amenaza la integridad física del sujeto. Luego, el sistema inmunitario debe poseer una capacidad de aprendizaje que le permita reconocer y defenderse contra nuevos antígenos. Debe poseer una *memoria* que conserve el recuerdo de las nuevas formas encontradas. El reconocimiento, el aprendizaje y la memoria forman parte de los procesos y mecanismos estudiados por las ciencias cognitivas. Comúnmente, estos modelos se relacionan con el cerebro, que sería su semejante biológico. Adelantamos que el sistema inmunitario es una *red cognitiva*, por el hecho de que en ambos casos, estamos en presencia de redes biológicas de propiedades similares (o al menos comparables), cuyos comportamientos cognitivos representan propiedades emergentes. Esta discusión sería inimaginable si no dispusiéramos de resultados de estudios recientes provenientes de la inmunobiología experimental, que ponen en evidencia las características de red del sistema, tal como lo explicaremos más adelante. Nuestro enfoque requiere, no sólo nuevas perspectivas epistemológicas, sino también cambios significativos en lo referente a la investigación inmunológica y a sus aplicaciones que no serán tratadas en este artículo.

De manera opuesta, las primeras teorías concebían "la molécula de anticuerpo como una especie de pegamento universal, capaz de interactuar con cualquier forma antigénica y adoptar su forma complementaria, de suprimir (o aislar) al antígeno y de recordar la configuración «aprendida»[8]". Aquí son pertinentes las comillas que enmarcan la palabra "aprendida". Como la naturaleza cognitiva del

[8] J. Urbain (1986), "Idiotypic networks: A noisy background or a breakthrough in immunological thinking?", *Ann,. Inst. Pasteur/ Immunologie* 137C: 57-64, p.56.

aprendizaje es inevitable, estamos obligados a referirnos a un modo cualquiera de aprendizaje. Pero, es evidente que, en este caso, el proceso implicado está lejos de ser cognitivo, puesto que, si lo fuera, podríamos deducir que el papel en el que aparece nuestra firma la ha "aprendido".

Tocamos ahí el núcleo de nuestro tema. En inmunología es común evocar la discriminación yo/ no-yo. Este enfoque deriva inevitablemente del hecho de que el sistema inmunitario manifiesta sus capacidades de discriminación al *interior* de un cuerpo. Este simple hecho tiene muchas consecuencias. Hasta hace poco, la inmunología ha seguido la tendencia adoptada en otras áreas de las ciencias cognitivas, que considera cualquier forma de capacidad cognitiva como una forma de tratamiento de la información. La información es introducida, y se supone que, seguidamente, el sistema actúa de manera apropiada sobre la información, de tal manera que produce una respuesta adaptada. Estas relaciones de entrada/salida, concebidas, en general, de acuerdo a la modalidad de programas internos para su tratamiento de "la información", constituyen la base de los enfoques *heterónomos* — los que consideran que el sistema está dirigido desde el exterior.

Este esquema heterónomo, fielmente adoptado por los inmunólogos, constituye la piedra angular de una concepción del cuerpo como pura exterioridad. Este cuerpo debe permanecer aislado de toda relación con el cuerpo primordial. El entrelazado no puede efectuarse a menos y hasta que la autoconstitución autónoma, origen de toda identidad biológica, ocupe el centro de la reflexión. En otros términos, hemos sido impulsados a reconocer que una dimensión esencial de la reinterpretación biológica del cuerpo debe cambiar: la autonomía propia de lo viviente es una condición necesaria para escuchar al cuerpo primordial[9]. Más tarde volveremos a hablar del tema. Por ahora, esforcémonos por definir paso a paso las etapas de este análisis.

[9] Cf. Las referencias en la nota 6; ver también: F. Varela, "Reflections on the circulation of concepts between a biology of cognition and systemic family therapy", *Fam. Process* 28: 15-24,1988.

De acuerdo a la concepción heterónoma, el antígeno se introduce en el sistema y la repuesta apropiada consiste en la producción de anticuerpos que llevan a la eliminación del antígeno. Pero, ¿qué determina la manera en que se forma el anticuerpo? A la inversa del sistema nervioso, el sistema inmunitario no cuenta con órganos sensoriales que puedan ser localizados en el espacio. Los anticuerpos circulan libremente en el interior del organismo y tienen más oportunidades de entrar en contacto con las moléculas que constituyen los tejidos propios del organismo (el "yo") que con los antígenos provenientes del exterior ("no-yo"). En resumen, el reconocimiento de los anticuerpos hace que sea inevitable que deba haber un medio para *reconocer lo que debe ser reconocido.*

Si el lector piensa que las cosas se vuelven demasiado complicadas, estamos de acuerdo: lo son. Sin embargo, es importante comprender bien que la necesidad de formular la manera de saber lo que es necesario saber hace que la definición heterónoma de la inmunidad como respuesta automática a cualquier cosa proveniente del exterior sea poco satisfactoria. La inmunología reciente (1950-1970) ha tratado de soslayar esta dificultad; no obstante, no ha conseguido mantener la integridad de la teoría heterónoma. Resulta natural que una solución más satisfactoria exija una revisión más radical de nuestra comprensión del sistema inmunitario como red autónoma. Pero procedamos etapa por etapa y examinemos con un poco más de precisión lo que "reconocimiento" puede significar en este contexto.

Decir que un anticuerpo "reconoce" un antígeno significa que se une a él químicamente y al hacerlo, lo neutraliza. Economía admirable cuando se trata de una molécula extraña y no de moléculas que son parte integral del organismo. Esta lógica simple es la base de la actitud de los inmunólogos que excluyen *a priori* la posibilidad de que los anticuerpos puedan unirse a las moléculas del yo sin desencadenar las consecuencias bien conocidas de las enfermedades autoinmunes. Fuera de esas situaciones patológicas, el organismo no desencadena reacciones inmunitarias deletéreas contra sus propios tejidos. Se trata del fenómeno de tolerancia identificada, en 1900, por Ehrlich con el nombre de *horror autotoxicus.*

Una idea importante, elaborada progresivamente, proponía que el reconocimiento de los antígenos desconocidos podía estar fundado, en gran medida, en la imprecisión de los mecanismos, es decir, en el hecho de que un anticuerpo puede unirse con distintos grados de afinidad a un amplio abanico de configuraciones moleculares. De esta manera, una gama de 10 variedades de anticuerpos basta para mantener vivo a un renacuajo, mientras que el hombre posee, para mantenerse vivo, un repertorio de más de 10 variedades de anticuerpos. En otras palabras, existen muchas maneras en que el sistema inmunitario está completo con el fin de asumir sus tareas. Este hecho hace que el concepto del yo/ no-yo sea más complicado.

Doble discriminación, doble reconocimiento: es necesario saber lo que es un antígeno "no perteneciente al yo" antes de poder reconocerlo como tal; es lógico que el reconocimiento de lo no perteneciente al yo implica el conocimiento del yo. Recordemos que el escollo estriba en el hecho de que el reconocimiento conlleva la destrucción. Predicamento diabólico que se podría resumir de la siguiente forma: "Las teorías clásicas requieren, por un lado, comparar lo que diferencia a las estructuras del yo de aquellas del no-yo, mientras que por otro lado, éstas imponen la ignorancia de la existencia del yo o la amenaza de una autodestrucción inmunológica[10]." Nos gustaría llamar a este predicado la *paradoja instauradora* de la inmunología clásica: no se puede defender sin reconocer, ni reconocer sin destruir.

Ahora podemos analizar la importante etapa siguiente del pensamiento inmunológico que enfrenta el problema evocado anteriormente, manteniendo la visión heterónoma del sistema inmunitario: la teoría de la selección clonal. Esta es el resultado de los trabajos de Niels Jerne y de Macfarlane Burnet durante los años 50 y ha dominado la inmunología hasta una época muy reciente, así como, durante el mismo período, la visión simbólica/ informática de la cognición ha dominado las ciencias cognitivas.

---

[10] A. Coutinho, L. Forni, D. Holmberg, F. Ivars, N Vaz (1984), "From an antigen-centered, clonal perspective on immune responses to an organism-centered network perspective of autonomous activity in a self-referential immune system", *Immunol.Revs.* 79: 151-168, p. 152.

La primera idea básica propuesta por Niels Jerne es la de un repertorio de anticuerpos que residiría en el cuerpo permanentemente. A la inversa de las teorías instruccionistas, él sugería que la producción de anticuerpos *precedía* y, en cierta forma, anticipaba la intrusión del antígeno. Esta teoría iba absolutamente en contra del sentido común de la época. Hoy sabemos que existen cerca de 10 anticuerpos diferentes, cada uno de los cuales puede unirse a una variada gama de perfiles moleculares ("*degeneracy of binding*"); actualmente no hay duda acerca de la noción de repertorio interno.

Queda por explicar cómo lo que en un principio es un conjunto accidental de anticuerpos puede ser configurado por una población dada de antígenos, es decir, por lo exterior. Sabemos que los encuentros antigénicos dejan rastros en el sistema: los anticuerpos que se les unen aumentan sustancialmente de cantidad. Esto constituye la clave de la *respuesta* inmunitaria. De hecho, este tipo de fenómeno nos obligará finalmente a tomar en cuenta las propiedades cognitivas del sistema, pero aún no llegamos a ese punto. Jerne ha propuesto una solución notable para eludir esta dificultad. Para esto, invoca a Darwin y la selección natural. Aunque el antígeno no funcione como la señal o la instrucción necesaria para la formación del anticuerpo, éste puede, no obstante, hacer una *selección* a partir de los anticuerpos existentes, unirse a ellos y hacer crecer su cantidad. Fue Macfarlane Burnet quien propuso un mecanismo específico mediante el cual este proceso selectivo podía ser interpretado en términos de un mecanismo linfocitario. La idea fundamental era que cada linfocito sólo puede exhibir (y producir) un solo tipo de anticuerpo, de tal manera que a cada antígeno se asociaría una subclase de familias de linfocitos o *clones*. El contacto entre el antígeno y el clon lleva a la proliferación de las células de ese clon que conduce entonces a una mayor producción de anticuerpos de ese tipo particular, neutralizando así al intruso antigénico (Burnet, 1959). De esta manera, el linfocito y la población de anticuerpos evolucionan bajo la presión selectiva del antígeno. No se trata de una determinación genética sino *antigénica*. Aún se utiliza la etiqueta de "determinante antigénico" (*antigenic determinant*).

La teoría de la selección clonal era una respuesta brillante al difícil problema de saber cómo procede el sistema inmunitario cuando debe enfrentar una situación nueva. Los problemas cognitivos aparecen aquí disfrazados de práctica evolucionista: se traspone una vasta escala temporal por otra más corta; y se traspone igualmente un ambiente exterior por uno interior orgánico.

La respuesta de la teoría de la selección clonal a la eterna pregunta de la autotolerancia era simple. Esta teoría formula la hipótesis de que el depósito básico de anticuerpos está, de hecho, incompleto: son *precisamente* los clones los que pueden reconocer las moléculas del yo que faltan. Solución simple que sólo traduce en otros términos los datos iniciales del problema, puesto que el enigma no puede ser resuelto *a priori* a través de la comprensión de simples mecanismos genéticos. Por lo tanto, no podemos invocar ni el proceso genético, ni el proceso selectivo en el adulto para efectuar este ajuste necesario. La única solución viable era atribuir esta etapa intermedia al embrión, en el que la teoría de la selección clonal suponía la supresión de los clones anti-yo. Esta idea se expresa habitualmente diciendo que el organismo "aprende" la discriminación yo/ no-yo durante la ontogenia.

Es así como los antiguos problemas cognitivos vuelven a aparecer por la ventana después de haber sido expulsados por la puerta. El modelo de selección ya no basta: aún estamos obligados a introducir un proceso de aprendizaje para delimitar un yo, aunque este proceso esté relegado a la vida embrionaria. ¿A través de qué mecanismos específicos podrá realizarse esto? Como dice Burnet, los clones anti-yo pueden ser evitados "suponiendo que en esa etapa del desarrollo embrionario el contacto antigénico lleva a la muerte de la célula[11]". Curioso desplazamiento de la lógica anterior: los mecanismos que aseguran la discriminación entre el yo y los antígenos durante la vida embrionaria participan de manera inversa en el adulto, en la discriminación entre los diferentes antígenos. En

[11] M. Burnet (1959), *The Clonal Selection Theory of Acquired Immunity,* Nashville: Vanderbilt U. Press, p. 58.

el segundo caso, el contacto antigénico lleva a la eliminación del antígeno. En el primer caso, el contacto antigénico lleva a la eliminación de la célula. La modificación teórica consiste aquí en separar los dos polos de la "paradoja instauradora" inmunitaria, atribuyendo un aspecto de la discriminación al adulto y el otro al embrión. Además, los clones que deben ser eliminados durante la vida embrionaria son elementos del yo: la autodestrucción se encuentra implícita al interior mismo del marco que intentaba evitarla. Está claro que el ajuste deficiente de estos dos procesos contradictorios es sólo una cuestión de coherencia epistemológica.

La teoría de la selección clonal ha sido un terreno fértil de inspiración en la práctica experimental. Esta ha llevado a Burnet a postular la posibilidad de engañar al sistema inmunitario introduciendo células en un embrión con el fin de hacerle tolerar moléculas que normalmente le son extranjeras. Esto ha probado ser experimentalmente correcto y ha sido claramente establecido que la tolerancia ha sido adquirida. Sin embargo, más tarde se ha demostrado que dicha tolerancia no es de propiedad exclusiva del embrión. En efecto, el adulto también puede poseerla; por lo tanto, el aprendizaje no puede ser limitado a un período específico de tiempo. Esto plantea interrogantes difíciles para la teoría de la selección clonal.

Más aún, la noción de un repertorio completo, que excluye a los determinantes del yo, ya es problemática. Esta exclusión tan inocente es de hecho demoníaca, si se recuerda la gran gama de perfiles moleculares a los que se une el anticuerpo, capacidad que se encuentra en los orígenes de la noción de completitud evocada anteriormente. La noción simple de eliminación clonal del anti-yo sólo es verdaderamente simple a partir de la hipótesis de una especificidad de los anticuerpos, idea importante para los pioneros de la inmunología y reforzada por las técnicas de vacunación que en escasas situaciones, inducen una clase estrecha de anticuerpos contra un patógeno. Sin embargo, resulta que, según las pruebas, la regla "un anticuerpo, un antígeno" es falsa. Si eliminháramos una cantidad suficiente de clones al punto de suprimir toda respuesta a las moléculas del yo, eso volvería a privar al organismo de la

posibilidad de responder a una cantidad enorme de antígenos potenciales. Así, la "completitud" *(completeness)* de su sistema estaría comprometida; la protección de la defensa inmunitaria estaría tan perforada como un colador. Así vemos que la fórmula de un repertorio completo del tipo "con la exclusión de", pensado para resolver el dilema de dos formas contradictorias de reconocimiento necesarias para la discriminación del yo/ no-yo, lleva a otra contradicción. Esta nueva contradicción deriva de la precisión necesaria para el reconocimiento del yo, precisión incompatible con la hipótesis de la completitud del sistema para responder al universo antigénico.

Como sucede a menudo en las ciencias, al final de algunas décadas de dominio, las flaquezas de la teoría de la selección clonal se han hecho cada vez más evidentes. No nos extenderemos más en los diferentes tipos de prueba experimental a este respecto. Fue, sin embargo, la combinación de esta insatisfacción teórica evocada anteriormente, junto con algunas observaciones empíricas lo que abrió, en los años 70, las nuevas perspectivas que ahora abordaremos.

2. La red inmunitaria como afirmación del yo

Los dilemas ya evocados siguen intactos a menos que se acepte abandonar el concepto de partida del *horror autotoxicus*. Actualmente está establecido que los anticuerpos que circulan normalmente se unen a muchas moléculas del yo (si no a todas...) tanto en el embrión como en el adulto. Ya no se puede interpretar que estos anticuerpos están dirigidos *contra* las moléculas del yo. De hecho, los anticuerpos que se encuentran en gran concentración al principio de las enfermedades autoinmunitarias son inofensivos en el índice normal.

Pero hay algo más importante aún: no se puede ignorar que los anticuerpos que están en circulación, a los que se les imputa la responsabilidad de la discriminación del yo/ no-yo, forman parte del yo. Esto implica la existencia de anticuerpos que se unen a otros anticuerpos, los que en jerga técnica se denominan anticuerpos "anti-idiotípicos". De esta manera, los anticuerpos libres y aquellos unidos

a las membranas celulares no son elementos individuales separados en clones, sino que están íntimamente unidos los unos a los otros para formar, hablando con propiedad, una organización en *red*. Esta idea fue propuesta por primera vez por Jerne (aún él) en 1974[12], e hicieron falta diez años para que fuera aceptada. Es así como, por primera vez, se vuelve progresivamente necesario considerar el sistema como funcionando según su propia dinámica interna[13]. Esta perspectiva pone en duda la noción de un sistema inmunitario heterónomo.

Imaginemos un antígeno extraño penetrando en el organismo. Una parte del antígeno, su determinante antigénico, será reconocida por ciertos anticuerpos. Designemos este perfil molecular con la letra E (por "epítope" en jerga técnica). De acuerdo al modelo antiguo, se trataría de un anticuerpo anti-E listo para eliminar el antígeno portador de E. El reconocimiento interviene sólo entre los dos y el antígeno conserva su papel selectivo. En la teoría de la red, este diálogo privado no tiene sentido. En primer lugar, porque ya no se trata de un vínculo bipolar, sino de un vínculo múltiple entre E y varios anticuerpos anti-E. Pero, de una manera más significativa, porque ahora es necesario tomar en cuenta los anticuerpos que se vinculan con los idiotopos de los anti-E. Estos últimos, a su vez, tendrán anticuerpos que se vincularán a sus idiotopos y así sucesivamente... De ello se deduce que siempre nos enfrentaremos a clases de anticuerpos que, al menos parcialmente, se parecerán al epítope E. Para decirlo de una manera más simple, el antígeno sólo podrá entrar en la red en la medida en que *ya* exista un anticuerpo en circulación que posea un perfil molecular suficientemente parecido al suyo, es decir, una "imagen interna". El antígeno deja de ser un "determinante" y se transforma en una perturbación leve en una red en movimiento perpetuo. Esto significa que los efectos

[12] N. Jerne (1974), "Towards a network theory of the immune system", *Ann. Immun. Inst. Pasteur* 125C:373-389, N. Jerne (1984), "Idiotypic networks and other preconceived ideas", *Immunol. Revs.* 79:5-24.

[13] I. Lundqvist, A. Coutinho, F. Varela y D. Holmberg (1989), "Evidence for the functional dynamics in an antibody network", *Proc. Natl. Acad. Sci* (EE.UU.).

de la intrusión de un antígeno variarán de acuerdo al conjunto del contexto de la red.

Ahora podemos comprender cómo el punto de vista heterónomo acerca del sistema es debilitado a través del simple examen de la lógica de la red que lo construye. Evidentemente, cuando el inmunólogo inyecta una gran cantidad de un antígeno, la respuesta inmunitaria se parece a una respuesta heterónoma. El concepto de red subraya el hecho de que ésta es una situación de laboratorio, muy artificial. Normalmente *no* recibimos grandes cantidades de un antígeno. Poseemos cantidades reducidas de diferentes moléculas del cuerpo que cambian con el tiempo, y estamos expuestos a cantidades reducidas de moléculas durante la respiración y la alimentación. En otras palabras, el sistema es fundamentalmente una unidad autónoma sensible a todo tipo de modulaciones que actúan cambiando sutilmente los niveles internos, pero no es en ningún caso una máquina para producir respuestas inmunitarias. Es así, por ejemplo, como los animales que no son expuestos a ningún antígeno después del nacimiento *(antigen-free animals)* desarrollan un sistema inmunitario absolutamente normal; esto está en clara contradicción con la teoría de la selección clonal que habría predicho la atrofia del sistema inmunitario en esas condiciones.

La siguiente etapa de nuestro análisis consiste en abandonar la noción de sistema inmunitario como instrumento de defensa destinado a responder a los sucesos exteriores y concebirlo en términos de afirmación de sí mismo, estableciendo una *identidad del cuerpo* en la unidad del ambiente somático a través de la especificación y la conservación de los componentes moleculares. En este nivel, el sistema inmunitario adquiere toda su dignidad y toma su lugar en la investigación actual de las redes biológico-cognitivas. En estas redes, la riqueza de las interconexiones genera niveles internos mediante procesos distribuidos. Más precisamente, la dinámica de encuentros entre anticuerpos y células regula la cantidad de linfocitos y los niveles de perfiles moleculares en circulación. Esta idea es análoga a la de la red de especies que otorga una identidad a los ecosistemas al interior de su medio ambiente. Lo que es interesante aquí es que esta ecología de linfocitos modifica el cuerpo

al interior del cual ella existe, cuerpo que, al mismo tiempo, la determina.

La danza mutua entre el sistema inmunitario y el cuerpo es la clave del punto de vista propuesto aquí, puesto que es esta danza la que permite al cuerpo poseer una identidad cambiante y plástica a lo largo de su existencia y de sus múltiples interacciones. El establecimiento de la identidad del sistema, en lugar de ser una reacción contra los antígenos, es una empresa *positiva* y creativa. Aquí se considera que el trabajo de precisar la identidad es, a la vez, lógico y biológicamente primario; la historia antigénica (ontogénica) modula la unidad en segundo lugar.

El lector se habrá seguramente dado cuenta de que esta presentación del sistema inmunitario es esquemática y simplificada: nuestro propósito aquí es definir algunos conceptos fundamentales y una arquitectura lógica. En este sentido es importante comprender bien lo que abarca hoy la "perspectiva de red" en inmunología. La realidad de los anticuerpos antiidiotípicos es muy reconocida actualmente y, para la mayoría de los inmunólogos, la red inmunitaria significa una cadena de antiidiotipos sucesivos. Sin embargo, toda la riqueza de los procesos de la red y de sus propiedades emergentes permanece sin ser completamente aclarada en inmunología. La cantidad de investigaciones experimentales que estudian específicamente los problemas de la red inmunitaria pueden contarse con los dedos de una mano, y en cuanto a las exploraciones teóricas, éstas apenas comienzan[14]. El progreso depende esencialmente de nuestra capacidad de abandonar la metáfora de la inmunidad como aparato de defensa (incluso cuando ésta es empleada a través de un proceso de red idiotípica). Se trata de aprender a considerar esta red como el artífice de una identidad molecular, es decir, verla como un *sistema inmunitario autónomo*. Por lo tanto, se pone énfasis en la fabricación de un yo autónomo en lugar de situarlo en la heteronomía y la exterioridad.

---

[14] F. Varela y A. Coutinho, "Immune networks: Getting on to real ground", *Immunol. Today* (por imprimir); Perelson, A. (Ed.), *Theoretical Immunology*, 2 vols. Addison Wesley, Nueva Jersey.

Tratemos ahora de comprender cómo se aclara esta concepción de una red autónoma con la "paradoja instauradora" de la inmunología, y el eterno problema de la discriminación entre el yo y el no-yo. La respuesta es de hecho muy simple y no difiere de aquella que conlleva la solución de cualquier paradoja: basta con salir del ámbito en que ésta es aún válida. En el presente caso, esto significa que el sistema inmunitario fundamentalmente *no discrimina* (no puede discriminar) entre el yo y lo que no pertenece al yo. La red, en movimiento perpetuo, sólo puede ser perturbada o modulada por intrusos antigénicos, respondiendo así sólo a lo que ya está y se le *parece*. De ello se deduce que todo antígeno que perturbe la red inmunitaria es por definición un "antígeno en el interior" y, en consecuencia, no hará sino modular la dinámica de la red. Cualquier cosa que no pudiera comportarse de esa forma sería simplemente un sinsentido y bien podría desencadenar una respuesta inmunitaria "refleja", es decir, una respuesta producida por procesos casi automáticos que sólo se encuentran en la periferia de la red. La antigua discriminación *yo/ no-yo* se transforma, en esta perspectiva, en una distinción *yo/ sin sentido.*

Normalmente, los antígenos se asimilan a través de la alimentación o la respiración y su presencia es regulada mediante los múltiples circuitos en los que pueden participar. Así se conservarán pocos índices de antígenos y de anticuerpos. Es también el caso de los elementos constitutivos del cuerpo. A lo largo del desarrollo, las moléculas del yo interactúan con los componentes inmunitarios de tal manera que sus niveles se conservan dentro de ciertos límites, puesto que estos están integrados dentro de una actividad inmunitaria permanente. Así, se puede demostrar, por ejemplo, que el nivel de la hormona renina es controlado por múltiples anticuerpos normalmente presentes en el sistema inmunitario del individuo. Se debe mencionar que este fenómeno permanente no es necesariamente (y generalmente no lo es) un problema de estabilidad: la variedad y el *"turnover"* de los componentes son demasiado elevados. Se trata más bien de un problema de viabilidad, o de trayectorias que cambian constantemente que, a pesar de todo, se aíslan en el interior de ciertos límites. Esto es lo que queremos

decir cuando hablamos de la afirmación *positiva* de una identidad molecular: lo que somos en el ámbito molecular y lo que nuestro sistema inmunitario es, representan dos entidades en relación recíproca de co-evolución.

El lector, acostumbrado a pensar en la inmunidad en términos de defensa, se impacienta. Usted bromea, dice; piense, por ejemplo que cuando nuestra inmunidad está debilitada, como ocurre en el caso del Sida, inmediatamente nos vemos asolados por agentes patógenos. Es indiscutible que el sistema *también* es capaz de dar una respuesta inmunitaria para reaccionar frente a una infección. Esta respuesta se producirá cuando la cantidad de antígenos sea demasiado importante, o cuando éstos penetren demasiado rápido y pongan en funcionamiento mecanismos específicos que acaban en una respuesta inmunitaria, inclusive en la inflamación en el lugar de la herida. Estos mecanismos son, en su mayoría, independientes de los procesos de red descritos anteriormente. Esta reactividad inmunitaria del tipo "reflejo" ha sido el objeto casi exclusivo de la inmunidad clásica. No se trata de negar el papel de defensa, sino de verlo como una manifestación o caso límite de algo más fundamental: una identidad celular y molecular históricamente construida.

3. Las identidades del cuerpo: la emergencia y la autonomía

Hemos realizado nuestra deconstrucción del yo inmunitario como defensa heterónoma, y su reinterpretación como afirmación del yo autónomo de una de las dimensiones de la identidad corporal. Ahora es necesario retroceder y referirnos a la orientación más fundamental evocada al principio de este artículo. Para hacerlo, debemos esclarecer algunas nociones planteadas en el transcurso de nuestra discusión sobre el caso específico de las redes inmunitarias. Dentro de toda una clase de modelos biológicos, las redes inmunitarias sólo constituyen un ejemplo entre otros, y, por lo tanto, es esencial comprender los desafíos del estudio de estas redes.

La característica esencial del enfoque de red es comenzar por elementos simples que se vinculan dinámicamente unos con otros.

Cada elemento sólo opera en su ambiente *local*, de tal manera que no existe un agente exterior para hacer funcionar el sistema. Pero, debido a la constitución en red, existe una cooperación *global* que surge espontáneamente cuando los estados de todos los "agentes" que participan alcanzan una condición universalmente satisfactoria. En este sistema no es necesario recurrir a una unidad de tratamiento central para guiar toda la operación. El paso desde las reglas locales a una coherencia global constituye el núcleo de lo que, durante los años cibernéticos, se acostumbraba llamar autoorganización. Hoy preferimos utilizar una terminología diferente: hablamos de propiedades *emergentes* o globales, de redes dinámicas o no lineales, de sistemas complejos[15].

No existe una teoría unificada para estas propiedades emergentes. Sin embargo, para nosotros, éstas representan una idea importante y radicalmente nueva para la ciencia moderna. Está claro que esas propiedades emergentes son útiles en todas las áreas —vórtices y láser, oscilaciones químicas, células, redes genéticas, esquemas de desarrollo, genética de poblaciones, redes inmunitarias y neuronales, ecología y geofísica—. Estas diversas áreas tienen esto en común: en cada caso una red da origen a propiedades nuevas. La labor de los investigadores es tratar de comprender estas propiedades en su generalidad.

En el ámbito de las propiedades emergentes, los fenómenos de lo viviente ocupan un lugar único debido a su modo específico de surgimiento que les permite autoindividualizarse: éstos afirman su identidad de una manera que sólo es propia a la vida desde sus orígenes. El interior y el exterior de los sistemas no vivientes son indiferentes el uno al otro: no es el caso de un organismo. Esto puede ser intuitivo para el biólogo, pero, sin embargo, el problema de la autonomía de lo viviente sólo ha sido abordado de manera correcta recientemente. Desde Claude Bernard, los biólogos suponen que existe un "medio interior" que es el producto del mismo

---

[15] Para más detalles ver P.Dumouchel y J.P. Dupuy, *L'auto-organisation: De la physique au politique,* Ed. du Seuil, París, 1983.

organismo. Sin embargo, la autonomía de lo viviente hunde sus raíces de manera más profunda. El ambiente interno está formado y determina los tejidos que envuelve. La red inmunitaria está formada y forma el ambiente somático. La célula viviente es, a la vez, el producto y la productora de su red de reacciones metabólicas; es decir, es autopoiética. Este cambio de formulación no expresa sólo una preferencia filosófica; refleja la necesidad de comprender los procesos de lo viviente no sobre la base de operaciones por entradas y salidas, sino a través de su *cierre operacional*[16]. Un sistema con cierre operacional es aquel en que los resultados de sus procesos son precisamente estos mismos procesos. (Mencionemos que la palabra "cierre" no significa cierre o ausencia de interacciones; hace más bien referencia a la naturaleza autorreferencial del proceso). Por lo tanto, la noción de cierre operacional es una forma de precisar la categoría de los procesos que, en el transcurso de su funcionamiento, dan vueltas sobre sí mismos, constituyendo así una red *autónoma*.

Aprehender los procesos de lo viviente como esta forma específica de emergencia marcada por la autonomía implica una comprensión radicalmente nueva de la relación entre el interior y el exterior. En la medida en que la circularidad del cierre no es considerada como eje central, la constitución del cuerpo permanece separada entre lo externo, que es dado por el exterior, y lo interno, que deriva del interior. En lugar de eso, la red autónoma propone una visión de la interioridad en que el interior y el exterior son términos que se *definen recíprocamente*, que están enredados mutuamente, puesto que ninguna determinación puede originarse de manera aislada, ni desde el interior ni desde el exterior. De acuerdo con la tradición que domina a la biología —la herencia de Claude Bernard— la autonomía orgánica se funda en la estabilidad del ambiente interno. Por consiguiente, la inmunidad debe ser vista como la defensa de esta estabilidad. Se oculta el hecho de que el organismo es la fuente misma de esta estabilidad.

---

[16] Para una discusión más detallada ver F. Varela, *Autonomie et Connaissance*, *op. cit.*

Las formaciones emergentes resultantes del cierre constituyen un modo procesal fundamental único para lo viviente. Estas nos obligan a revisar el aspecto complementario de cada identidad que es su modo de conocer. El punto de vista tradicional propone que las relaciones cognitivas determinen las representaciones del ambiente, tal como lo hemos visto en nuestro análisis de la inmunología. Si afirmamos que la función de estos procesos es representar un ambiente independiente, estamos obligados a considerarlos como pertenecientes a la clase de sistemas fundamentalmente heterónomos, en que la información debe existir previamente, y constituye el impulso para su cognición. Este impulso constituye "las premisas iniciales" a partir de las cuales el sistema elabora un comportamiento, es decir una respuesta. Pero ¿cómo podemos distinguir el impulso de la reacción en un sistema autogestionado y altamente cooperativo como el cerebro? Existe, desde luego, un intercambio de energía y de materia, pero ¿dónde termina la "información" y dónde comienza el "comportamiento"? En un texto reciente, Marvin Minsky ha estudiado el problema y vale la pena citar sus observaciones *in extenso*:

"¿Por qué estos procesos son tan difíciles de clasificar? Antiguamente podíamos juzgar las máquinas y los procesos a través de la forma en que éstos trasformaban los materiales brutos en productos terminados. Pero es insensato hablar de los cerebros como si fabricaran pensamientos del mismo modo en que las fábricas hacen automóviles. La diferencia radica en que los cerebros utilizan procesos que se modifican a sí mismos — y eso significa que no podemos separar estos procesos de los productos que ellos manufacturan. El cerebro elabora específicamente recuerdos que modifican la forma en que pensaremos posteriormente. *Entre las actividades principales del cerebro encontramos la de realizar cambios en su interior.* Debido a que la idea de los procesos automodificables es nueva en relación con nuestra experiencia, aún no podemos confiar en nuestra sensatez cuando tratamos estos temas[17]. "

---

[17] M. Minsky, *The Society of Mind*, Simon and Schuster, Nueva York, 1987, p. 288. Las cursivas son nuestras.

Esta cita es notable debido a la ausencia de toda noción de representación. Minsky no dice que la actividad principal del cerebro es representar el mundo exterior; dice que ésta consiste en realizar automodificaciones permanentemente. ¿Adónde se fue la noción de representación? En efecto, está a punto de efectuarse un viraje importante en el mundo de las ciencias cognitivas bajo la influencia de sus propias investigaciones. Este viraje exige que abandonemos la idea de que el mundo es independiente y extrínseco para aproximarnos a la idea de que el mundo es inseparable de la organización por cierre, propio de lo viviente. Los procesos cognitivos no funcionan a través de la representación: en lugar de *representar* un mundo independiente, *hacen surgir* un mundo como un ámbito de distinciones inseparables de la estructura misma del sistema cognitivo[18]. Esto constituye el aspecto complementario de la circularidad en la determinación entre el interior y el exterior de una identidad corporal.

No podemos extendernos aquí y referirnos a las formas más concretas que estas ideas toman en el ámbito de la biología contemporánea. El debate acerca de la disciplina inmunológica podrá servir como ejemplo e incentivar la imaginación del lector. Otras dimensiones importantes para ser consideradas serían la organización celular[19], la forma y el desarrollo multicelular[20], los fenómenos neurocognitivos[21] y la comunicación[22].

El cuerpo es el lugar de intersección de las diferentes identidades que surgen a partir del cierre, que hace que el interior y el exterior sean confusos. *Somos* y *habitamos* dicho entrelazado; nuestro cuerpo no posee una identidad externa única, sino que constituye

[18] F. Varela, *Connaître: Les sciences cognitives*, Seuil, París, 1988.

[19] L. Margulis, *Symbiosis and Cell Evolution*, Freeman, San Francisco, 1982.

[20] L. Buss, *The Ontogeny of the Individual*, Princeton Univ. Press, Princeton, 1988.

[21] H. Maturana y F. Varela, *The Tree of Knowledge, op. cit.*

[22] T. Winograd y F. Flores, *Understanding Computers and Cognition*, Addison Wesley, New Jersey, 1987.

una red dividida y embrollada sin otro fundamento sólido que su propia determinación procesal. Esto nos sirve de base para volver al cuerpo primordial que expresa un "interior no subjetivo":

> *"El organismo en el cual cada suceso es una causa y un efecto de todos los demás y, en este sentido, causa de sí mismo, plantea el problema de la autoproducción de la unidad, o más precisamente, de una totalidad... que trabaja sobre los materiales que le son propios, por decirlo así, emanan de sí mismos*[23]*."*

4. El cuerpo evocador: la falta en la raíz

Así hemos hecho "el círculo completo" y volvemos a nuestra problemática del comienzo: la intimidad del habitáculo de nuestra propia carne. Hemos mostrado bastante de nuestra orientación para formular una conclusión preliminar que orientará el trabajo posterior. Para hacer esto, proponemos hacer la distinción entre *tres cuerpos*.

En primer lugar, el cuerpo fundamental o *primordial*, tal como se ha descrito aquí al referirse a los procesos a través de los cuales el cuerpo se constituye como unidad. Se trata, una vez más, del cuerpo biológico visto como la intersección de surgimientos biológicos. Este cuerpo, anterior al yo y por lo tanto preontológico, lleva implícita una comprensión previa del significado del ser.

El segundo cuerpo, es el cuerpo *formativo*, la imagen del cuerpo en vías de constitución, tal como se reconoce en el *infans* y el niño. Este cuerpo actualiza la articulación entre el cuerpo primordial y la imagen del cuerpo constituido durante el proceso de su advenimiento. Es un cuerpo que nace en el curso de los primeros años de vida y cuyo redescubrimiento a través de una práctica de transformación exige cierto trabajo, puesto que está encubierto por los efectos de la expulsión, oculto en una corporeidad postural y fisiológica.

[23] M. Merleau-Ponty, *Leçons des Cours* , PUF, París, p. 71.

En tercer lugar, tenemos la imagen del cuerpo *constituido*, el *"ego-self body"*, la identidad determinada social y lingüísticamente. Este cuerpo es coextensible al cuerpo propio en la aceptación fenomenológica[24], y dispone de una conciencia más o menos amplia de su condición dividida y sin fundamento.

Estos tres cuerpos están siempre presentes en nuestra experiencia y forman un entrelazado "vertical". Nuestra hipótesis es que habitamos un cuerpo que es la reconstitución continua de sus identidades emergentes. Pero el movimiento que define esta condición siempre es animado por la falta de algo: la identidad no se encuentra en ninguna parte, salvo en su condición autónoma, en sus propios procesos circulares que se autoafirman. Inevitablemente, este modo autónomo siempre está cerca del desmoronamiento. Si la vida consiste siempre en una actividad en relación con lo que falta, ésta también es su impulso, su deseo de continuar.

Visto de esta manera, el *"ego-self body"*, el sujeto separado y que desea, en términos psicoanalíticos, está ***hundido hasta la raíz misma de la vida***, y es inseparable del cuerpo primordial. Se encuentra aquí, con toda su fuerza, la condición de posibilidad del cuerpo primordial anunciada al principio de nuestro artículo y que se ha designado como "el entrelazado". Recordemos que esto quiere decir que existe una circulación entre los dos polos de acercamiento del cuerpo primordial: una óptica que lo ve como proceso biológico, y una praxis que se dirige a la experiencia del sujeto que habita el cuerpo. La *topología* definida por la interdeterminación de estos dos polos es la única en que el cuerpo primordial puede ocurrir y ser considerado en toda su riqueza, sin separarlo ya sea de las raíces de su intimidad o de su vida.

La exploración del entrelazado es aún un proyecto fragmentario donde queda casi todo por hacer. Sin embargo, se pueden esbozar algunos puntos de reflexión referentes al psicoanálisis, como práctica de transformación. La praxis del psicoanálisis, que también se puede

---

[24] M. Merleau-Ponty, *Phénoménologie de la perception*, Gallimard, París, 1962.

denominar una "erotología"[25], nos enseña la irreductibilidad del deseo a un principio común (principio del placer), ya sea la felicidad, el lenguaje o la ciencia[26]. El *cuerpo que desea* o *erógeno*, modalidad de existencia del cuerpo accesible al psicoanálisis, no se define en relación con las categorías de la "psique" y de lo orgánico como exterioridad biológica. Se podría decir que pertenece, precisamente, al entrelazado.

Es muy sorprendente que a menudo la teorización analítica coloque al cuerpo erógeno en una topología diferente: una posición de exterioridad interior, como el inconsciente. Eso nos remite a su manera de enfocar la idea de lo "irrepresentable". Se podría decir que el pensamiento analítico por excelencia *tiende hacia* lo impensable, lo irrepresentable: la conceptualización fundadora freudiana del síntoma histérico es el ejemplo más evidente del papel central de lo no representacional en el proyecto psicoanalítico. La revisión de la identidad inmunitaria como ejemplo de surgimiento autónomo es, por cierto, un modelo discursivo. Pero esta óptica no es antinómica con un acceso al cuerpo primordial dotado de una modalidad de "saber" prediscursivo. La metáfora del cuerpo "cognitivo", en lugar de fortalecer las categorías ontológicas del sujeto y del objeto, las deconstruye. El "saber" del sistema autónomo y la noción de la codefinición que deriva de él modifican en profundidad la idea de la pareja de opuestos interior /exterior y descentran, de una vez, nuestra manera de ver lo que es propiamente interior y exterior. En la perspectiva de la autonomía de lo viviente, la maraña de los tres cuerpos designa la falta en el origen mismo de la identidad de lo viviente. Es común que esta intuición se presente metafóricamente como "principio de lo viviente". Aquí se trata de su despliegue en un discurso científico que proporciona otra topología para seguir la elaboración de los fenómenos que escapan a la representabilidad del cuerpo que desea y la alienación del deseo en el lenguaje. ¿Cómo van al encuentro de este cuerpo primordial el discurso y la práctica analítica?

---

[25] J. Lacan, "Le Séminaire", *L'Angoisse*, 1962-1963 (inédito).

[26] J. Rajchman, *Le savoir-faire avec l'inconscient*, William Blake and Co. 1986.

Reflexionemos acerca de aquello que edifica nuestro estatus de sujeto pensante. El pensamiento se funda en la noción de representación, pero en su raíz está el deseo. "El pensamiento sólo es un substituto de un deseo alucinador... sólo el deseo puede impulsar a nuestro aparato síquico a trabajar[27]." Al originarse en el deseo, el pensamiento funciona a costa del rechazo del mismo deseo que lo impulsa. De cambio en cambio, la función misma del pensamiento discursivo es de no bastar para rendir cuentas de ese deseo que se representa como falta, pérdida necesaria para la posibilidad misma de la representación. El que el pensamiento, la conciencia discursiva, sea *necesariamente*, una amenaza (quizás la palabra es demasiado fuerte) para el deseo, y produzca un distanciamiento del cuerpo primordial preontológico, deriva de la constitución del sujeto a partir de su división. La ciencia, la cultura se edifican en la sombra, la "sutura"[28] de esta línea divisoria constitutiva.

Por naturaleza, el pensamiento analítico está condenado a alejarse de una cierta modalidad de existencia del cuerpo, y el proyecto en que se inscribe este texto ("la evocación del Ser a través del cuerpo"), puede ser comprendido como un recuerdo de esta debilidad. En efecto, la sexualidad como objeto conceptualizado por el pensamiento psicoanalítico, nació de esta incapacidad constitutiva, inherente al pensamiento. Tomada, como la describe Bersani, en la historia, en la narración freudiana del desarrollo del cuerpo erógeno, la idea de la sexualidad compromete o rechaza un modo de funcionamiento que desea. En el texto freudiano, la sexualidad "no es un placer hipotético puramente corporal, sino que, desde un principio, es una señal del fracaso del pensamiento en dar cuenta de, en encontrar la expresión adecuada para la experiencia del cuerpo. El intento del pensamiento de replicar sus propios fracasos funcionales sería un fantasma[29]".

---

[27] S. Freud, *L'interprétation des rêves*, PUF, París, 1967, p. 482.

[28] Dor, *op.cit.*

[29] L. Bersani, *The Freudian Body*, Columbia Univ. Press, 1986, p. 65.

Pensamos aquí en el problema de la pulsión y, más precisamente, en su dimensión de fenómeno *afectivo*. La teoría "clásica" del afecto se apoya en gran medida en la diferenciación freudiana entre afecto y pulsión. La pulsión se compone de dos fenómenos, la carga afectiva y la huella mnésica: estos dos fenómenos no son idénticos aa la pulsión, pero la "representan". El afecto corresponde a un proceso de descarga cuya "manifestación final" es percibida como un sentimiento. Este se compone, a su vez, de dos elementos, el fenómeno de descarga motor o secretor y el fenómeno "subjetivo" que es la emoción[30].

En la práctica, el afecto se remite a su componente "representable", dotado de un estatus protopático, y se subordina de manera jerárquica a la discursividad. Por lo tanto, se relega a un estatus de "prerrealidad", de "aún no representado". Sin embargo, tanto la teoría[31] como la práctica nos enseñan que el afecto no es algo que se aplique *en* el discurso, una *coloración* separable, de cierta manera "centrifugable" del contenido discursivo. Nos parece que considerar el afecto, ya sea como fenómeno de descarga, o bien como fenómeno de marcación, limita su tratamiento psicoanalítico. ¿Estos límites se deben a que el afecto sólo sería accesible a partir de su dimensión discursiva o representacional?

Finalmente, parecería que el afecto se presenta como el lugar privilegiado en que se manifiestan los dos polos del entrelazado. Quizás, éste constituye una vía de acceso privilegiada al cuerpo primordial. En este punto, nuestra discusión sugiere una *intrincación* entre el registro discursivo y afectivo para debilitar la jerarquía habitual y poner los dos registros en posición de cocreación[32]. Cualquiera que sea el destino de esta hipótesis, la noción

---

[30] Freud, *Métapsychologie*, Gallimard, París, 1985.

[31] Pensamos en los trabajos de D.W. Winnicott acerca del estatus de la pulsión en el espacio de transición, y en el trabajo de P. Aulagnier acerca del pictograma y "lo originario".

[32] De manera más particular en el contexto de este artículo, en que la inmunidad cumple un papel central, los fenómenos psicosomáticos son de una importancia ejemplar como dominio de elaboración ulterior.

de la interioridad tal como es revisada aquí remite a todo un abanico de problemas. Sólo podemos evocar uno de ellos: el problema de las raíces corporales y discursivas del afecto que plantea interrogantes acerca del estatus del afecto y acerca de lo que éste significa para el psicoanálisis (y para cualquier otra práctica de transformación) de la "cara íntima" del afecto que permanece fuera de la escena.

II

MENTE

II

MENTE

1. Cognición

FORMAS VIVAS DE DAR SENTIDO A LAS COSAS

Una vía intermedia para la neurociencia

Este texto fue presentado en uno de los primeros simposios convocados para reflexionar sobre las implicaciones de la complejidad y los sistemas auto-organizados que dominaron la discusión científica en los años 80. Para mí fue la ocasión de trazar un bosquejo inicial de las grandes líneas de mi visión acerca del origen del significado y del conocer en los seres vivos. Es lo que llamo aquí una «vía intermedia para la neurociencia», y que posteriormente tomaría más forma bajo el concepto de enacción. *("¿De dónde viene el origen del significado perceptual?").*

Este simposio se convino bajo la idea de que, en años recientes, muchos científicos han puesto en tela de juicio los conceptos tradicionales de desorden y orden. Este es un pensamiento ambiguo y nada modesto. De hecho, esta conferencia es una crítica explícita y desenfadada al programa[1] que predomina entre los principales pensadores relacionados con las ciencias de la mente y el comportamiento, junto con la proposición de una alternativa. Sin embargo, nos parece claro el hecho de que podemos presentar estas ideas únicamente en la medida en que se hayan producido los cambios correspondientes en las demás disciplinas científicas que hacen posible esta reunión.

(Varela, F. (1984), Living ways of sense-making: A middle path for neuroscience, in: P. Livingstone (Ed.), *Order and Disorder: Proceedings of the Stanford International Symposium*, Anma Libri, Stanford, pp.208-224).

[1] Este término se usa en el sentido técnico que le da Lakatos (1970) en su análisis de la actividad científica como una acción colectiva.

El orden es el orden, en relación con alguna persona o ser vivo que adopte tal postura frente al mismo. En el mundo de los vivos, el orden es de hecho inseparable de las formas en que los sistemas vivos hacen sentido, de tal manera que pueda decirse de ellos que tienen un mundo. Asociamos tal actividad con animales dotados de sistemas nerviosos y consideramos a dichas acciones como actos cognitivos. Es evidente que el orden y el conocimiento están entrelazados. Como científicos, buscamos los mecanismos que nos permitan revelar el origen de tales actos cognitivos, los cuales dan cuenta incluso de nuestro sentido del orden. Sin una referencia a este hecho básico fundamental, el objetivo de este simposio no podría lograrse. Nuestra postura es que deben reevaluarse los actuales puntos de vista respecto del modo en que los seres vivos dan sentido a las cosas. Es a esto a lo que nos referiremos ahora a través de un ejemplo ilustrativo que usaremos con el fin de apoyar el argumento general.

1. El color

El mundo del color alrededor nuestro parece ser tan naturalmente un elemento constitutivo de los objetos que una fotografía a color nos parece análoga a lo que debe estar sucediendo en las mentes de aquellos que poseen una visión normal de los colores. Si algo parece ser más blanco que otra cosa, debe ser porque estamos recibiendo más luz de él, y si algo se ve rojo, debe ser porque estamos recibiendo luces de una onda espectral de una longitud particular.

Sin embargo, si examinamos la situación con mayor detenimiento, recibiremos una sorpresa. Una gran sorpresa, de hecho, porque si efectivamente medimos la luz que emana el mundo a nuestro alrededor, veremos que no existe una relación de uno a uno entre los flujos de luz en distintas longitudes de onda y los colores asociados con los objetos[2]. Esto se revela de manera trivial si pensamos en los grandes esfuerzos que realizamos como fotógrafos para ajustar una cámara, incluso para leves variaciones en la iluminación.

[2] Para una discusión sobre este tema, ver Beck (1972) y la demostración particularmente clara en Land y McCann (1971).

En este punto pensamos: "Bueno, debe ser que el ojo tiene un dispositivo inteligente de compensar los cambios en la iluminación y extrae de todos los cambios alguna cualidad estable de los objetos tales como su reflectancia." De hecho, el razonamiento físico dice que podemos pensar de la luz como un elemento que perturba la retina como el resultado del reflejo proveniente de un objeto bajo una iluminación dada, y por ende, como proporcional al producto de ambos factores. Es dable, entonces, a partir de este razonamiento, que el mismo flujo de luz puede ser el resultado de un número infinito de combinaciones de reflectancias e iluminaciones.

Quedamos así en una situación enteramente circular. La retina es perturbada por un flujo de luz, el producto de la iluminación y la reflectancia. Sin embargo, ni la iluminación ni la reflectancia son constantes en nuestro campo visual normal, y ninguno de los dos factores es conocido con anticipación. Con el fin de "recuperar" la reflectancia como la propiedad de un objeto, sería necesario que hiciéramos algunas comparaciones con un elemento estándar de referencia. Ciertamente no nos dedicamos a hacer este tipo de comparaciones cada vez que observamos la naranja sobre la mesa, por muy desigual que sea la iluminación, porque tal comparación requeriría una iluminación fija como referencia. Entonces, nos vemos forzados a concluir que la sensación del color es independiente del flujo y distribución de longitud de onda de la luz reflejada por un objeto.

Esto es, en sí, algo perturbador para nuestro tren de pensamiento usual, pero las cosas son peor de lo que parecen. No sólo se viola nuestra ingenua suposición relativa a la correlación entre la longitud de onda y el color, sino que se viola, además, otra suposición igualmente ingenua: que la experiencia del color debe asociarse con una propiedad local del objeto coloreado en lugar de la del campo visual en su totalidad.

En este punto, usted puede estar preguntándose de qué estamos hablando, de tal manera que permítanme referirme a una ilustración. Consideremos dos proyectores colocados de la manera que se indica en la Figura 1. Uno de los proyectores está equipado con un filtro

rojo. Ahora colocamos una mano frente al proyector sin filtro y, dado que no estamos haciendo nada más que obstruir el paso de la luz blanca en esa región, esperamos ver una sombra roja en un fondo rosado. Eso es precisamente lo que vemos. Sin embargo, si ahora proyectamos una sombra desde el proyector rojo, esperaríamos ver una mano blanquecina en un fondo rosado. Pero la mano se contrasta fuertemente en un tono azul-verdoso tan convincente como puede serlo. Sin embargo, el espectrómetro nos dice que el flujo de luz de esa región azul-verdosa es blanca en la composición espectral.

Este nítido experimento de la llamada "sombra de color"[3] se destaca como uno de los mejores ejemplos que conocemos de una anomalía, para usar una terminología de Kuhn. Nos hemos referido a él para que cumpla precisamente esa función. Como es el caso de todas las anomalías, ha sido tratada como una paradoja. La pregunta que normalmente se hace es: "¿Entonces, cuál es el color *real* de la sombra?" o bien se afirma: "Esta es una ilusión cromática". Como si pudiéramos todavía basarnos en la correlación color-longitud de onda, lo cual ciertamente no podemos hacer, como acabamos de ver. Más aún, lo que queda tan claro en este experimento es, de hecho, algo que es normal en el funcionamiento del sistema visual, léase, la función de los contornos o las discontinuidades en el nivel de actividad en las neuronas retinianas y sus interconexiones, lo que hace de todo el campo visual una unidad.

Pensemos acerca de esto sólo un poco más con una demostración adicional que va un poco más allá de la demostración precedente. Consideremos los mismos dos proyectores de la Figura 1, pero ahora coloquemos a cada uno copias idénticas de una diapositiva que contenga una figura de tablero de ajedrez basado en grises, blancos y negros como se muestra en la Figura 2. Las dos diapositivas son superimpuestas de tal manera que estén alineadas en forma exacta. Ya que hemos mantenido el filtro rojo en uno de los proyectores, el patrón general es de una variedad de rosados en diferentes grados de saturación.

---

[3] Un fenómeno descrito por primera vez por Otto von Guericke, en 1672.

Ahora giremos en noventa grados una de las diapositivas. El resultado es una imagen completamente policromática que contiene pequeños cuadrados de color amarillo, azul y verde, así como rojos y rosados[4].

El efecto de este experimento es bastante elocuente: la policromaticidad surge cuando uno esperaría solamente diversos tonos de rosado. Este efecto cromático se debe a una simple desigualdad en las proporciones de blanco a blanco y rojo a rojo que se producen a lo largo de los bordes de los pequeños cuadrados lograda mediante la rotación de uno de las diapositivas. ¿Cómo puede ocurrir esto? De hecho puede concebirse que el flujo de luz que llega a la retina perturba a tres mosaicos de conos diferentes pero entremezclados que constituyen tres superficies retinianas: las retinas "roja", "verde" y "azul", cada una caracterizada por el fotopigmento de sus conos y con curvas de absorción superpuestas cuyas diferentes sensibilidades llegan a su punto más alto en alrededor de 590, 540 y 450 *nm* respectivamente.

Más aún, estas tres retinas no son de ninguna manera idénticas u homogéneas: la retina "roja" tiene una densidad de conos alrededor de cinco veces la de la retina "azul" y ligeramente inferior a la de la retina "verde". Además, como resultado de la conectividad interna de la retina, en el caso de cualquier flujo de luz que golpee el locus de la retina, el efecto que tiene la diferencia de la actividad local de las tres superficies de recepción depende de lo que sucede a cada una de ellas en el resto de la retina.[5] De este modo, se generan valores de referencia interna, y las desviaciones abruptas de tales valores de referencia en los niveles locales de actividad pasan a ser la diferencia que hace la diferencia; dentro de los límites de tales desviaciones de una referencia interna, se percibe un color uniforme.

---

4 Este experimento pertenece al tipo de fenómeno estudiado por E. Land (1964). El uso de rotaciones de tableros cuadriculados de grises como se presenta aquí fue descrito inicialmente en Maturana *et al.* (1968).

5 Una de las formas que puede adoptar es un uso masivo de interacción lateral. Esto se expresa de manera más precisa como un proceso cooperativo en el que cada valor local depende de valores en puntos colindantes. Véase Horn (1974) y Marr (1974).

De hecho, la cualidad inasible que llamamos "color" no puede explicarse si buscamos explicarla en términos de las propiedades de los objetos que han de ser mirados por el observador. En lugar de eso, podemos dar cuenta de la experiencia del color revelando su constitución a través de la coherencia interna de la actividad neuronal relativa que surge en el sistema nervioso.

Hay más detalles en la especificación de estas dimensiones cromáticas del comportamiento, pero no abordaremos estos tecnicismos.

2. El programa representacionista

En este caso, lo que nos resulta más interesante en esta ilustración acerca del color es el hecho de preguntarnos ¿Por qué debemos sentir que nos encontramos ante algo sorprendente? ¿Por qué pensamos que estamos obligados a describir las sombras de color como ilusorias y nos sentimos intranquilos por no encontrar una correspondencia entre longitud de onda y color?

La sorpresa e intranquilidad sólo pueden venir de una fuente: un hábito inveterado de nuestro pensamiento, tanto en las ciencias como en otras áreas. Queremos cuestionar este hábito inveterado, como se manifiesta explícitamente en los estudiantes de la neurociencia. Lo llamamos el *programa representacionista*. Tiene sus raíces en un complejo de ideas que se tornaron visibles en su actual forma con el surgimiento de la ingeniería de las comunicaciones y la ciencia computacional, pero que es una herencia mucho más antigua en las preferencias anglosajonas. "El sistema nervioso captura características del entorno y construye representaciones del mundo en el que vive el animal con el fin de utilizarlas como acciones de adaptación del organismo". Este programa considera a las características del medio ambiente que van a representarse tan poderosas y centrales como para constituir las pautas principales para el estudio de las formas y comportamiento neural.

Un análisis representacionista comienza aislando ya sea una modalidad de actividad particular de un componente neural (digamos,

una respuesta cromática en una célula ganglionar retiniana) o su desempeño en relación con un comportamiento particular (digamos, discriminación de colores), convirtiéndolos en una única característica. Luego, la característica aislada se explica en términos de alguna forma de *adequatio rei ad cerebrum* (adecuación de las cosas en el cerebro), ya sea a través de una descripción funcional a través de la cual se asume que encarna el objeto por su referencia a su supuesto valor biológico (digamos, el reconocimiento de los colores es adaptativo para los encuentros sexuales) o bien a través del análisis de los componentes neurales en términos de como supuestamente contiene la característica ambiental (la manera en que se codifica y procesa el rojo luego de que golpea sobre las superficies de recepción).

Una modalidad extrema de este punto de vista es proporcionada por la llamada doctrina de la célula única de la percepción sensorial[6], en la que la correspondencia cerebro-mundo es mediatizada a través de una actividad única de las células. Esta teoría ha sido criticada sistemáticamente en años recientes; sin embargo, los ítemes conceptuales que se proponen en la actualidad en su reemplazo, por ejemplo los canales sintonizados en base a frecuencias[7] son igualmente representativos. La actitud fundamental es todavía la misma: existe un mundo exterior que debemos conocer y la tarea del sistema nervioso es hacerlo mediante la aprehensión de sus características.

Los neurocientistas con seguridad alegarán que estamos distorsionando sus puntos de vista. ¿Acaso no es aceptada la noción de que existe un rol para modificaciones internas tales como el control eficiente de la percepción sensorial?¿Acaso no es ampliamente

---

[6] Ver Barlow (1972). Las siguientes citas de libros conocidos también dan el sabor fundamental. "El cerebro es un conjunto de células que no descansan, que continuamente recibe información, la elabora, la percibe y toma decisiones" (Kuffler y Nicholls, 1976, página 3). "He encontrado necesario el suponer que el ente que percibe tiene ciertas estructuras cognitivas que funcionan para recoger la información que ofrece el medio ambiente" (Neisser, 1976).

[7] Ver Richards (1979) para una discusión reciente.

utilizada la idea de los generadores centrales de patrones? O, dicho en términos generales (y más cerca de la versión del psicólogo cognitivo) ¿no es claro acaso el hecho que uno necesita alguna forma de programa interno para la integración sensorial/motora?

El punto al que queremos llegar aquí es, sin embargo, algo diferente. Las personas frecuentemente evitan una confrontación directa con una teoría rival no a través de la negación del punto del oponente, sino evitando su lógica interna y estética apremiante. En términos concretos, esto implica la siguiente movida interesante: reconocer el punto de vista rival, pero circunscribir su dominio de acción a un ámbito tan específico que no tenga ninguna importancia real en la explicación que buscamos. En el caso de la eferencia sensorial, digamos, el abandono se produce aproximadamente de la siguiente manera: Importará solamente en aquellas circunstancias en las que exista una necesidad de ajustar la sensibilidad del animal a un nuevo nivel de estímulos del medio ambiente; por lo tanto, es irrelevante la mayor parte del tiempo.

La aceptación de alternativas en principio no necesariamente significa que van a llegar al laboratorio. La mayoría de los neurocientistas está dispuesta a aceptar que no todo es una representación. Sin embargo, enfrentados con un comportamiento o pieza de sinaptología, buscan actividades que correspondan a un comportamiento adaptativo en el mundo. Este es un hábito conceptual muy arraigado alrededor del cual se ha dado forma a toda la neurociencia moderna. Es el hábito de tomar la descripción de las interacciones que experimenta el organismo como la pauta principal en la búsqueda de una explicación para sus consecuencias. Equivale a tratar a cada interacción como *instructiva* de la misma manera, de tal manera que sus consecuencias deben ser representaciones sobre las cuales opera el sistema nervioso para generar un comportamiento adaptativo. Y el hábito parece completamente reforzado por su éxito en el diseño de dispositivos de ingeniería y computadores. Pero sería una actitud simplemente descuidada de nuestra parte si usáramos tal estrategia en nuestros intentos de comprender lo vivo y el entendimiento sin mas examen.

Una forma de esta falta de cuidado es simplemente el no tomar en serio la arquitectura del sistema nervioso y lo que implica para su operación. Por ejemplo, por cada fibra nerviosa que ingresa al córtex que proviene de una célula ganglio-retiniana (vía el núcleo geniculado), existen alrededor de cien otras fibras de otras áreas cortical y subcortical que ingresan en la misma ubicación topográfica[8]. ¿Tenemos alguna esperanza de encontrar en las llamadas "señales" del mundo llevadas por el nervio óptico una guía acerca de lo que ocurrirá en el cortex?

A decir verdad, durante varios años se pensó que el descubrimiento de lo que entonces se llamó "detectores de características" en la retina y el cortex anunciaban un gran avance en nuestra comprensión de los mecanismos del cerebro porque eran la encarnación de las representaciones[9].

Sin embargo, muchos años después podemos ver que las respuestas celulares repetibles pueden obtenerse solamente bajo condiciones experimentales estabilizadas[10] y gran parte de la actividad neural observada parece extremadamente variable porque está influenciada por muchos factores diferentes. Por lo demás, pueden proporcionarse sistemas clasificatorios alternativos sin que ello implique descripciones de características y únicamente en términos de la coherencia interna de los parámetros locales, como la geometría del campo de recepción, trayectoria temporal de la respuesta u otras características particulares de la neurona[11].

Además, muchos estudiosos de la neurociencia están comenzando a darse cuenta de la necesidad de confrontar, en toda

[8] Ver, por ejemplo, la discusión en Braitenberg (1977).

[9] Maturana *et al.* (1959); Hubel y Wiesel (1962).

[10] Ver, por ejemplo, Horn and Hill (1969) para el efecto de la inclinación del cuerpo sobre la respuesta del cortex visual; Fishman y Michael (1973) y Morell (1972) para el efecto de la estimulación auditiva; y Hubel y Wiesel (1962), y Hoepner (1974) para leer acerca de otros parámetros que afectan las respuestas visuales.

[11] Rowe y Stone (1980) explican estos puntos de manera muy clara.

su tremenda dificultad, el hecho de que el análisis de célula única no nos llevará muy lejos.

De hecho, podemos describir la lógica de las operaciones del cerebro como una entreverada jerarquía de comparaciones de superficie a superficie. Recién ahora estamos empezando a tener alguna idea de cómo expresar esto en términos más explícitos en algunos casos particulares, tales como los mecanismos visuales[12], o cómo elaborar un mapa de tales actividades globales usando métodos ópticos que involucren a todo el cerebro.

Tomada en serio, esta tendencia de la investigación en la neurociencia implica un cambio radical en nuestros puntos de vista acerca del funcionamiento del cerebro. Para nosotros significa dejar de ver al cerebro como un dispositivo que recibe información en la forma de información sobre la cual actuar. Más bien, significa moverse hacia una visualización del cerebro como un sistema caracterizado no por sus entradas sino por el cierre operacional de sus dinámicas de estados definidas como un equilibrio relativo de la actividad entre superficies neurales, de tal manera que cada cambio de estado en el sistema pueda conducir solamente a otro cambio de estado en el sistema mismo. En este programa, la investigación interesante es hacer explícitas a las dinámicas y estabilidades de las diversas formas adoptadas por dicho cierre de estados internos bajo continuas perturbaciones ambientales[13].

En este punto, algunos de ustedes dirán: "Seguramente que esto es una broma suya; ¿usted está diciendo que lo que ocurre "afuera" no tiene importancia para el sistema nervioso, el cual se comporta como una solitaria mónada que da forma al mundo a su antojo?" Por supuesto que no estamos diciendo eso. El tratar al sistema nervioso como una mónada desconectada sería oscurecer la idea que estamos tratando de validar llevándola al absurdo. Llevar la idea al absurdo

---

[12] Ver, por ejemplo, Schwartzz (1980), Julesz (1981) y Marr y Poggio (1979) para algunas muestras de este tipo de trabajo, aunque sus autores no necesariamente estarían de acuerdo con nuestra interpretación. Ver además Maturana, Varela y Frenk (1972).

[13] Para ver la presentación original de este punto de vista, leer Maturana y Varela (1980).

significa transformarla en el opuesto lógico del paradigma representacionista. De hecho, para el programa representacionista es implícito (o explícito) que el mundo y sus cualidades son primarias y que informan directamente al sistema nervioso en términos de que recoger y sobre que actuar para la generación de un comportamiento adaptativo. Desde este punto de vista, las interacciones entre el organismo y su ambiente son de carácter instructivo. El opuesto lógico sería decir: "Es el organismo lo que reinventa al mundo y lo coloca en el futuro que desea —una especie de cerebro kantiano o una mente solipsista".

Nosotros apelamos en favor de una vía dialéctica intermedia. Deseamos recorrer el filo de navaja que hay entre el Scila del representacionismo y el Caribdis del solipsismo, plantándonos firmemente al medio[14].

---

[14] Claramente un mapa de tres puntos de la situación es excesivamente simplista y muchos encontrarán aceptables partes de cada enfoque. Esto es inevitable, pero se trata de buscar una clarificación presentando lo que, pensamos, son los aspectos teóricos clave que debemos seguir o rechazar como pautas para la investigación. Un compromiso tibio solamente puede producir resultados también tibios. He aquí una tipología de aproximaciones a la neurociencia que, muy convenientemente, tienen este enfoque en el centro:

a) *Percepción directa*

El programa de Gibson (1980) es el mejor exponente de esta forma inflexible de representacionismo. Junto a él, uno podría colocar la doctrina de la célula única de Barlow (1972) de la percepción sensorial.

b) *Computacional/representacional*

Esta sería la posición adoptada por el mayor número de psicólogos cognitivos, investigadores de la inteligencia artificial y neurobiólogos. Las principales pautas para la investigación se encuentran en alguna modalidad de ***adequatio rei ad machina***, pero en la que nos vemos forzados a ver un nivel "algorítmico" o "computacional", que es comprensible únicamente en términos de la propia estructura del sistema. Esto resulta más claro en la visión, por ejemplo, en el trabajo de E. Land y D. Marr. Para obtener mayores informaciones respecto a este tema, ver Fodor (1980) y Ullman (1980).

c) *Constructivismo*

En esta posición, uno ya ha cruzado y se encuentra al otro lado de la línea de el medio, y la clave es una estructura innata de guía interna, que es la más fundamental. Se hace poca o ninguna referencia a las limitaciones e interacciones que el observador describe en el ambiente. Ver, por ejemplo, von Glasersfeld (1979).

d) *Solipsismo completo a la Berkeley*

Según sé, no ha servido recientemente como inspiración para una estrategia de investigación.

¿Qué significa esto en términos pragmáticos? Significa tomar explícitamente en consideración las condiciones de nuestras descripciones, en virtud de las cuales vamos y venimos entre el organismo como sistema en su propia lógica interna y el organismo como una unidad en cuanto a sus interacciones. En este ir y venir entre estos dos dominios de descripción independientes nosotros, en nuestra calidad de observadores, establecemos como válidas ciertas restricciones fundamentales en el ambiente que el organismo debe satisfacer. Así, para la visión, al actuar como físicos, estamos de acuerdo en que no puede haber un comportamiento como la visión sin las interacciones con la luz. Y dicha luz no puede ser de longitudes de onda superiores a la del rojo ya que, en caso contrario, no podrá desviar electrones de órbitas moleculares. Tampoco puede ser de longitudes de onda mucho más cortas que la del violeta, porque rompería vínculos químicos. Pero dentro de estas restricciones amplias, los procesos que la luz gatilla al perturbar los receptores visuales pueden seguir una u otra de una gama completa de posibilidades. Cual se siga efectivamente está determinado en cada organismo de acuerdo con la estructura de su sistema nervioso y por ende, de acuerdo con la historia del organismo[15].

Para nosotros, el hecho de mirar a la coherencia internamente generada del cierre es el hilo de Ariadna que conduce al interior del problema. De este modo, la discriminacion del color como un comportamiento no existiría sin las interacciones con la luz como las hemos descrito; pero el color no se encuentra en las longitudes de onda de los flujos de luz. Esa propiedad de la luz que llamamos "longitud de onda" es ocasión de múltiples

---

[15] No debe escapar a la atención del lector en este punto que describimos cosas como la luz y la sinapsis a través de nuestro sistema nervioso y nuestra visión, y que cada descripción de las propiedades del organismo y su entorno se basa en operaciones de distinción que *nosotros* realizamos. De este modo, estamos en una situación circular de la que no podemos escapar, una condición constitutiva que es problemática solamente si tratamos de negarla, como lo hace la actitud representacionista. Esta enredada jerarquía de descripciones es una parte inseparable de lo que proponemos como la vía intermedia para la neurociencia, pero respecto de la cual no tenemos oportunidad de detenernos a reflexionar aquí. Para mayores informaciones sobre este tema, ver las discusiones en Maturana y Varela (1980) y Varela (1979).

modalidades de dimensionamiento de regularidades internas en todo el reino animal, todas ellas viables dentro de las restricciones amplias fijadas por lo que describimos como "propiedades de la luz", aunque no son color. El patrón efectivo de las coherencias en el sistema nervioso de un animal bajo perturbaciones visuales especifica su comportamiento cromático como un dominio de las distinciones cromáticas. Pero tal comportamiento cromático no es un calce óptimo con un ambiente ideal de objetos coloreados sino, en lugar de eso, es una de muchas posible vías de coherencia interna dentro de las restricciones generales impuestas por las estructuras primarias: el organismo y la luz. Nuestro universo de colores no es ni arbitrario ni la expresión de un calce óptimo con un medio ambiente dado. Se trata de la expresión de nuestras dinámicas de estados actuales como resultado de una de las posibles vías filogénicas viables dentro de muchas otras realizadas en la historia evolutiva de los sistemas vivos.

Por lo tanto, el origen del conocimiento (y del darle sentido a las cosas) no se parece al diseño de un sistema que ha sido optimizado para coincidir con un estándar exterior dado. Diríamos que se asemeja más bien a un remiendo, a una escultura dinámica, una construcción de estructuras a partir de los materiales disponibles para un organismo que ensambla a medida que aparecen en un flujo que sigue uno de muchos caminos posibles. En este proceso, cualquier vía de remiendo surgirá del proceso de *deriva natural*. La clave para este proceso es que las consecuencias de cualquier interacción han de encontrarse, no en la naturaleza de la perturbación que las gatilló, sino en la manera en que la estructura compensa dichas interacciones de acuerdo con su paisaje dinámico; el resultado general es el cambio en la estructura en un mantenimiento continuo de la integridad del sistema en su medium[16]. Una vía de deriva natural es una historia de las validaciones internas de las interacciones por parte de una entidad autónoma, como, por ejemplo, un organismo dotado de un sistema nervioso. Es,

[16] Usamos la expresión "remiendo" como la propone Jacob (1978) en el caso del pensamiento evolutivo. El término "deriva natural" fue acuñado por Maturana y Varela (1984).

una vez más, sólo recientemente que hemos comenzado a tener idea de cómo describir este tipo de fenómeno en una manera no-representativa, por ejemplo, como un paisaje dinámico. El término "creod", acuñado por Waddington en el contexto de la embriología, es posiblemente adecuado para designar a tales vías resultantes de distintas historias de deriva natural[17].

De hecho, la idea de que el sistema nervioso deba caracterizarse por tener un cierre operacional exige una revisión de la postura tradicional y se apoya en un cambio desde el énfasis desde el organismo como involucrado en una relación instructiva con su entorno a uno que ponga énfasis en la modalidad de autonomía propia del organismo en su entorno. De hecho, la importancia de la autonomía de lo vivo, expresada en el contexto científico moderno, es la clave para este cambio de postura[18].

4. Conexiones

Sería algo perturbador el hablar de todo esto y no encontrar proposiciones convergentes surgidas en años recientes desde muchas áreas diferentes de las ciencia, así como entre los pensadores sociales y filósofos[19]. En este sentido, nuestra descripción del sistema nervioso como caracterizado por el cierre operacional y la noción complementaria de la "autopoiesis", que describe la organización de lo vivo, son parte de una tendencia general de sensibilidades contemporáneas en la ciencia.

---

[17] El término "creod" proviene de las palabras griegas *chreia* "necesidad" y *odos* "un camino". Según Waddington, un creod es una "trayectoria de tiempo de cambio de desarrollo [surgida] de las características de la organización causal circular cerrada del sistema de genes y citoplasma... Los creods son un tipo de fenómeno que ocurre en muchas otros campos también". (Waddington, 1960, p. 82).

[18] Para una completa discusión acerca de la noción de autonomía y sus muchas ramificaciones, ver Varela (1979).

[19] Es particularmente interesante el hecho de que algunos filósofos anglosajones contemporáneos hayan recientemente abordado la crítica de la representación y su rol en las ciencias. Ver Rorty (1979) y Malcolm (1977).

De particular importancia para nosotros son las críticas actuales al programa dominante en la teoría evolutiva, con su énfasis en la adaptación como un proceso que conduce a un estado físico óptimo y la selección natural como el único mecanismo. Este programa adaptacionista está, irónicamente, en una línea diferente a la del pensamiento pluralístico darwiniano original que proponía a la selección natural como un mecanismo que solamente daba forma a las condiciones generales amplias para la viabilidad de un organismo. Dentro del rango permitido por la selección natural, las vías y los patrones específicos que toma un árbol filático no son generados por un proceso de optimización[20] sino que de hecho "dependen más crucialmente de las restricciones (embriológicas y fisiológicas) que restringen a posibles caminos y modalidades de cambio, de tal manera que las restricciones mismas se convierten, con mucho, en el aspecto más interesante de la evolución"[21]. Para el pensamiento evolucionista, el adaptacionismo adopta una función isomórfica al representacionismo.

De hecho, vemos que el hilo unificador que motiva gran parte del trabajo presentado en este simposio es la tendencia a remitirse al estudio de las coherencias internas generadas por la unidad como la pauta para comprender su fenomenología. Esto abarca muchos distintos ámbitos de descripción y muchos enfoques en los que la autorreferencialidad comienza a hacer su aparición, no como una molestia de la cual debe huirse, sino como el material básico a partir del cual puede entenderse la emergencia del orden y sentido, ya sea en la vida, entendimiento, textos o sociedad. Naturalmente, cada disciplina debe encontrar su propia especificidad bajo este amplio paraguas de resonancias denominada como "autoorganización".

Hagamos un resumen antes de concluir. El núcleo de la vía intermedia que proponemos para la neurociencia es el entrelazamiento (es decir, el elemento facilitador) entre las coherencias

---

20 Para una discusión explícita de este importante tema, ver Oster y Rocklin (1979).

21 Gould y Lewontin (1979, página 594). Este artículo es una excelente crítica del programa adaptacionista por parte de dos de los principales estudiosos de la evolución. El presente artículo debe mucho a la influencia del texto de Gould-Lewontin.

internas del cierre operacional del sistema nervioso y el resultado de su deriva natural. Estas nociones básicas son aplicables y apropiadas para los sistemas vivos en todas sus formas.

Pero algunas de estos nuevos puntos de vista acerca del orden y el desorden que nos han reunido aquí, y ciertamente la actual propuesta, encuentran grandes resistencias. Creemos que esto es así porque, tomado en serio, el programa propuesto aquí implica abandonar las nociones de representaciones e información como pautas útiles para la investigación. Las raíces de estas resistencias son profundas. Se trata ni más ni menos de las resistencias por abandonar edificios de la tradición de la física, en la que la principal pauta para la investigación era la causación como un agente iniciador, típico de la mecánica clásica. Se trata precisamente de la misma lógica de representaciones e interacciones instructivas. En una visión intermedia de lo vivo, la principal pauta de investigación es el elemento facilitador que hay entre la lógica interna de la unidad y su historia de selección por consecuencias. Esta es una situación enteramente nueva, tan accesible al estudio científico como la causación mecánica, pero crucialmente diferente y apropiada para un ámbito diferente de fenómenos[22].

Un punto final. Una cosa es considerar, y posiblemente estar convencidos, de que necesitamos abandonar una actitud representacionista para la comprensión del sistema nervioso de los animales y otra cosa muy distinta es aceptar con toda seriedad que lo mismo debería aplicarse a nuestra propia cognición, y la forma en que comprendemos nuestro propio conocimiento.

Esta aplicación de segundo orden de la vía intermedia, por darle un nombre, puede resultar incluso más difícil de absorber, ya que involucra el abandono del hábito inveterado de confiar en el mundo como un punto de referencia físico que debe ser representado o bien de tomar al sujeto como agente iniciador y por tanto también como punto de referencia. Si tenemos razón,

---

[22] Para una discusión acerca de la mecánica clásica en este contexto, ver Prigogine y Stengers (1979).

nuestra vida humana, nuestra experiencia presente, no es sino uno de los muchos posibles credos de conocimiento, en donde el inmenso trasfondo de nuestra estructura biológica y prácticas sociales es inseparable de la regularidad que discernimos tanto en el mundo como en el ser. Cuando llevamos esta lógica a sus últimas consecuencias, podemos entender al mundo en el cual nos encontramos como algo que no está ni aparte ni distante, aunque también podemos entenderlo como un mundo en el cual ya no quedan puntos fijos de referencia[23].

[23] En la tradición occidental, esta epistemología de lo sin-fundamentos tiene su expresión en pensadores como Merleau-Ponty y Heidegger, pero se expresa en un contexto incomparablemente más rico en la tradición *madhyamika* de la India medieval.

LA VISIÓN DEL COLOR
Un caso de estudio acerca de las bases de la ciencia cognitiva

Los colores, tan bellos y aparentes en el mundo, resultan estar preñados de una inseparable simbiosis entre mundo y organismo, que es precisamente lo que la visión enactiva del conocimiento proclama. No es sorpresa, pues, que la visión de los colores me haya ocupado como tema de investigación por muchos años, ya que resulta un ejemplo encarnado y concreto de lo que, de otra manera, puede aparecer un tanto abstracto y general.

> *«[El color] es "el lugar donde nuestro cerebro y el universo se unen", dice [Cézanne] ... No se trata entonces de colores "simulacro de los colores de la naturaleza", se trata de la dimensión de color, aquella que crea de si misma a si misma identidades, diferencias, una textura, una materialidad, un cualquier cosa»*[1].

Esta afirmación evoca al color como una dimensión, un lugar que no existe con anterioridad a quienes se reúnen allí; se produce *durante* la reunión misma. Nuestro propósito en este escrito es abordar esta cuestión sobre la base de nuestro trabajo experimental *comparativo* acerca de la visión del color, esto es, el estudio de la visión del color en diversas especies animales. Nuestra intención no es proporcionar una visión científica actualizada de este campo, sino considerarlo como un caso de estudio que revela cuestiones fundamentales en la ciencia cognitiva.

(Este texto ha sido condensado de un artículo que aparecerá en *The Behavioral and Brain Sciences*. Su contenido le debe mucho a un escrito de Evan Thompson, presentado al Departamento de Filosofía, Universidad de McGill, Canadá, en diciembre de 1988.)

[1] M. Merleau-Ponty, *L'Oeil et l'esprit* París, Gallimard, 1964, página 67.

Este escrito se desarrolla en tres etapas. En la primera parte, revisamos brevemente algunos trabajos actuales sobre el estudio de la visión del color. Esta visión será luego llevada a un límite crítico en la segunda etapa, a través de lo que nos gusta denominar el argumento comparativo. Este pretende demostrar la modalidad en la cual la visión de color es una actividad inserta ecológicamente en lugar de una forma de procesamiento de información. Advertimos de inmediato al lector que *no* interpretamos esto de ninguna manera como una forma de punto de vista subjetivista en el sentido de que el color es un tipo de sensación, ni como una visión lockeana en el sentido de que el color es una forma de disposición secundaria. El argumento comparativo, sostenemos, permite ir más allá de esas posturas clásicas. Esto se hace en la parte tercera y final, en la cual exponemos un punto de vista enactivo del color.

1. Ontología actual del color

Tal como se le enseña a todo niño en edad escolar, la ciencia moderna del color debe su origen a Newton, quien afirmó que el color es una sensación producida por las longitudes de onda que constituyen la luz[2]. Sin embargo, contrariamente a este punto de vista heredado, *no existe* una correspondencia uno a uno entre el color percibido de una superficie y la composición espectral e intensidad de la luz reflejada desde dicha superficie. Esta independencia relativa del color percibido de la composición espectral ha sido extensamente documentada en personas en la actualidad.

Esta independencia se manifiesta de manera más clara en dos fenómenos complementarios. En el primero, el color percibido permanece relativamente estable a pesar de grandes cambios en la iluminación, fenómeno que se conoce como *constancia* del color. En el segundo fenómeno, la *inducción* del color, dos superficies que reflejan una luz de la misma composición espectral pueden percibirse

[2] I. Newton, *Optics*, Nueva York, Dover, 1952 (basado en la edición de 1730), páginas 124-125.

como si tuvieran colores distintos dependiendo del entorno en el cual se los coloca[3]. Dados estos dos fenómenos decisivos (y una cantidad de evidencia concurrente), es simplemente fútil asimilar el color a la composición espectral. Tales intentos simplemente reflejarían que uno no comprende el fenómeno en absoluto, ya que la ciencia del color actualmente se basa en la distinción misma entre color y el contenido espectral de la luz[4]. De hecho, para poder establecer que un organismo tiene visión del color, uno prueba cuáles *combinaciones* de longitudes de onda gatillan *la misma* percepción del color.

¿Existe actualmente algún candidato para reemplazar a la longitud de onda en aras de una reducción objetivista del color? Algunos han propuesto recientemente que el color debería identificarse con la *reflectancia* espectral de la superficie[5]. La reflectancia se define como la proporción en cada longitud de onda entre el incidente y la luz reflejada. Por lo tanto, es una propiedad relativamente estable, ligada a los objetos y puede especificarse sin una referencia detallada a la microestructura de las superficies. Esto, a su vez, haría posible fenómenos tales como la constancia y la inducción. La postura neoobjetivista alega que el color es una propiedad objetiva del mundo y que puede identificarse con la reflectancia. El color, como lo percibimos, por el contrario, sería indeterminado con respecto al color objetivo ya que nuestra percepción no especifica completamente las reflectancias de las superficies y solamente nos da "tipos de colores definidos

---

[3] Para otros ejemplos de constancia en los humanos, ver Edwin Land, "The Retinex Theory of Color Vision", *Scientific American*, t. 237, 1977, páginas 108-128 y en un contexto comparativo, D. Ingle, "The Goldfish as a Retinex Animal", *Science*, t. 225, 1985, páginas 651 a 653; V. Budnik, H. Maturana, F. Varela, "Chromatic Induction: A Comparative Study", 1990.

[4] Para leer un excelente análisis filosófico de este tema, ver J. Westphal, *Color: Some Philosophical Problems from Wittgenstein*, Oxford, Basil Blackwell, 1987, página 75.

[5] Esto se explica con mayor claridad en R. Hilbert, *Color and Color Perception: A Study in Anthropocentric Realism*, Stanford, Center for the Study of Language and Information, 1987. También en P. Churchland, "Reduction, Qualia and the Direct Inspection of Brainstates", *Journal of Philosophy*, t 82, 1985, página 8-28; M. Matthen, "Biological Functions and Perception Content", *ibid*, t. 85, 1988, p 5-27.

antropocéntricamente y no los colores mismos"[6].

Esta última frase debe entenderse en el contexto de las actuales investigaciones acerca de la visión del color y —más precisamente— en el contexto de lo que se conoce como visión computacional del color, en la que el problema central es precisamente proponer algoritmos explícitos y redes neurales que permitan a un dispositivo biológico o artificial *determinar* la reflectancia superficial con un conocimiento incompleto, esto es, solamente con el conocimiento del contenido espectral de la luz reflejada por las superficies, que es todo lo que la retina recibe[7]. No es necesario que entremos en detalles acerca de esta línea de trabajo. Baste decir que es un típico caso de un problema inverso mal planteado. En esencia, estos son los tres elementos por considerar:

1) *Baja Dimensionalidad*: Uno observa que las reflectancias pueden describirse como yacientes dentro de un espacio de pocas dimensiones: solamente algunas funciones básicas son suficientes para abarcar el espacio apropiado. Asimismo, uno asume que existen menos parámetros que describan la reflectancia que parámetros que describan la clase de fotorreceptores que toman muestras de cada punto en la imagen. Por lo tanto, la visión de color computacional está restringida fundamentalmente por la baja dimensionalidad tanto de los tipos de estímulo como de receptores. Específicamente, los humanos tienen tres fotorreceptores, y las reflectancias naturales necesitan de tres a seis funciones base[8].

2) *Computaciones Globales*: La actividad local de un fotorreceptor en sí misma no es significativa. Lo que es relevante es

[6] R. Hilbert, *op. cit. supra* n.5, página 27.

[7] Para una buena introducción técnica, ver L. Maloney, *Computational Approaches to Color Constancy*, Applied Psychology Lab. Stanford University, Informe Técnico 1985-01, 1985 y A. Hulbert, *Color Computation in the Visual System*, A.I. Memorandum N° 814, Cambridge, MIT, 1984, y "Formal Connection between Lightness Algorithms", *Journal of the Optical Society of America A*, t. 3, 1987, páginas 1684-1693.

[8] L. Maloney, *op.cit.supra* n. 7, página 60.

la interacción global en el largo alcance, que transforma la luminescencia en "luminosidad", un nivel de actividad más cercano a la reflectancia. Existen varios de estos "*algoritmos de luminosidad*" equivalentes[9]. Todos pueden entenderse como una manifestación —en el sistema viviente— de las interacciones laterales y circuitos de reingreso que son típicos tanto de la retina como del sistema visual, conducentes a valores especificados internamente en lugar de los valores sensoriales crudos.

3) *Segmentación*: Incluso bajo computaciones globales de baja dimensión y al nivel de las redes, las reflectancias son aún indeterminadas. Un elemento clave que falta es la manera en que una escena se segmenta en los segmentos relevantes sobre los cuales se realizará el cálculo de la reflectancia. Así, deben aplicarse algunas suposiciones adicionales acerca de las superficies (lo abrupto del cambio, distribución de valores promedio, y así sucesivamente).

La unión de estos tres elementos permite regularizar el problema mal planteado de recuperar la reflectancia con diversos grados de éxito (o fracaso) y con diversos grados de verosimilitud biológica. No pretendemos ocuparnos aquí de esos detalles.

¿Qué significa esto en términos de la actual ontología de los colores? El objetivista diría que esta subdeterminación es precisamente la base de sus puntos de vista, porque existen diferencias en la reflectancia que no podemos detectar y, por lo tanto, solamente tenemos el punto de vista del ojo humano del color objetivo. Sin embargo, esto parece ignorar un punto importante que fuera revelado por los resultados de la visión computacional del color, y que implican una circularidad intrínseca en el argumento. Para conocer la reflectancia, debemos especificar la dimensionalidad (número de clases de receptores) y la segmentación superficial de la escena. Pero el objetivista asume que el color (reflectancia *qua*) es una propiedad intrínseca que ayuda a explicar como el sistema

---

[9] A. Hulbert, *op. Cit. supra,* n. 7.

visual efectivamente segmenta una escena, y por que la evolución produjo un número dado de clases de receptores. En otras palabras, los tres elementos que están involucrados en el cálculo de la reflectancia son mutuamente inter-dependientes y no tienen una jerarquía de precedencia lógica.

Esto ya deja en aprietos a la ontología del color objetivista. Volveremos a esto en la última Sección. Mientras tanto, hay más y ahora me referiré a lo que llamo el argumento comparativo.

2. El punto de vista comparativo

En lo que acabo de decir acerca de la visión computacional del color, es fácil perder de vista el fenómeno del color en un contexto más amplio. De hecho, existe una ciencia del color de más alto nivel que comprende elementos de la química, física, fisiología, ecología, sicología e inteligencia artificial. Deberíamos recordarnos a nosotros mismos lo que el color *es* dentro de estos puntos de vista de más alto nivel. La partícula existencial representa aquí los componentes necesarios de su fenomenología experiencia, a la cual debemos dar precedencia.

La definición más común de color es la descripción tridimensional de tono, saturación y luminosidad, que, en conjunto, definen un *espacio de color*. Consideremos ahora al tono. Los tonos pueden ser ya sea únicos o binarios. Un tono único es aquel que es puro, en el sentido de que no contiene otros componentes cromáticos: azul, verde... Los tonos únicos pueden ser opuestos: azul y amarillo, verde y rojo. Un tono binario, por contraste, es uno que sí contiene otros componentes cromáticos. Algunos tonos son necesariamente binarios, tales como el naranja, mezcla de rojo y amarillo. Los tonos binarios están siempre ubicados entre dos tintes únicos en el espacio de color.

Esta estructura de tonos únicos/binarios es, en sí misma, un síntoma de que algo falla en la visión objetivista del color, ya que no encontramos nada en la reflectancia que tenga esta estructura. La respuesta es obvia, es decir, que esta estructura única/binaria

puede ser meramente una estructura de la experiencia del color, pero no del color objetivo. Pero esta respuesta coloca el peso de la prueba en el lado objetivista, ya que *prima facie* el color es tono, saturación y luminosidad como propiedades necesarias. Decir que se trata solamente de propiedades de la experiencia, equivale a afirmar que los colores son propiedades de los objetos, pero que el rojo, azul, amarillo o azul, no lo son. Tales problemas indican que la reducción objetivista ha perdido de vista la dimensión fenoménica.

Llevemos el argumento un paso más allá. Como mencionamos antes, la visión de color humana es tricromática, ya que puede ser representada en un espacio con tres variables independientes. Las variables más adecuadas son, contrariamente a lo que se cree comúnmente, no las curvas de sensibilidad de los fotopigmentos de los conos retinianos, sino una combinación de ellos a los que se refiere como *canales oponentes cromáticos*. Estos son bien conocidos para los fisiólogos y los psico-físicos y, por lo general, son una luminancia no-oponente (o canal acromático), un rojo menos verde (o canal tritanóptico), y un amarillo menos azul (o canal deuteranóptico)[10]. Estos canales constituyen los ejes de lo que nos referiremos como un *dominio cromático*. Un dominio cromático debe distinguirse de un espacio de color, ya que tiene, como sus dimensiones, el tono, la saturación y la luminosidad. Estas dimensiones especifican lo que el color es y por lo tanto, especifican los fenómenos del color a su propio nivel. Un dominio cromático, por otro lado, tiene como sus dimensiones a los canales de color funcionalmente especificados de un perceptor dado, los cuales especifican el color al nivel de su encarnación.

Podemos ahora formular claramente el argumento comparativo en pasos sucesivos como sigue:

---

[10] Para una presentación del punto de vista clásico, ver L. Hurvich, *Color Vision*, Sunderland, Sinauer, 1981. Para una derivación teórica de los canales, ver G. Buschbaum, A. Gottschalk, "Trichromacy, Opponent Colours Coding and Optimum Information Transmission in the Retina", *Proceedings of the Royal Society of London* t. 220, 1983, p. 89-113.

(a) Un dominio cromático determina un espacio de color;

(b) Dado que los dominios cromáticos dicen relación con la encarnación en una clase de receptor dada, lo mismo sucede con el espacio de color.

(a) Ya resulta obvio el cómo pueden explicarse los tonos binarios y complementarios apelando a la tricromaticidad de nuestra visión de color. Podemos esperar algo similar para las otras dimensiones del espacio del color. De hecho, la dimensión de luminosidad corresponde al eje del canal acromático, la dimensión de tono a la periferia máxima en el espacio abarcado por los canales, y la saturación a los puntos dentro de los contornos de los tonos. Esta correspondencia proporciona un puente entre las propiedades fisiológicas de la visión del color y las características fenomenológicas del espacio del color. Por lo tanto, sugiere cómo nuestro dominio cromático determina el espacio del color.

(b) El estudio de la visión del color ha sido tradicionalmente muy antropocentrista. Sin embargo, existe una creciente conciencia en el sentido de que la visión del color es prevalente en la mayoría de los vertebrados y en muchos sistemas invertebrados y que "la verdadera culminación de la evolución de la visión del color en los vertebrados probablemente se de en los animales diurnos altamente evolucionados, los cuales posiblemente estén mejor representados por aves diurnas y es dentro de estas especies donde deberíamos buscar una visión del color significativamente más compleja que la nuestra..."[11] De hecho existe firme evidencia de que algunas aves diurnas, tales como la paloma y el pato son al menos tetracromatos e incluso

[11] J. Bowmakerm "Color Vision and the Role of Oil Droplets", *Trends in Neuroscience*, t. 3, 1980, páginas 41-43.

pentacromatos[12]. También hay firme evidencia de la presencia de tetracromaticidad en peces (tales como el pez dorado y la breca japonesa)[13]. El sistema visual de aquellas especies diversas parece tener cuatro (quizás cinco) canales, en contraste con tres en los seres humanos, lo cual conduce a dominios cromáticos de cuatro o cinco dimensiones[14].

Muchas personas, cuando escuchan esta evidencia, responden preguntando: "¿Bueno, dónde están los colores adicionales que un tetracromato o pentacromato ve?" La pregunta es entendible pero ingenua. Un tetracromato, por ejemplo, no puede ser imaginado como un ente que hace una distinción más fina entre, digamos, los tonos rojos y amarillos. Tal habilidad constituiría un incremento en la *resolución* dentro de nuestro dominio cromático. El ser un tetracromato significa que el espacio del color tiene una *dimensión* enteramente nueva. ¿Que podría esto significar?

Un *Gedankenexperiment* posible sería imaginar qué pasa cuando cambiamos tonos únicos a partir de, digamos, el color rojo. En un sistema tricromático, dando un círculo de tonos (una dimensión menos que el espacio del color), un rojo puede solamente moverse hacia ser más amarillo o más violeta, a lo largo de los dos ejes de una línea. Pero en un sistema tetracromático, los lugares de los tonos se convierten en una superficie; desde un punto identificado con un tono puro, podemos movernos en infinitas direcciones mientras seguimos permaneciendo dentro de la superficie de los tonos puros. Por ende, no hay manera de delinear

---

[12] S. Jane, J. Bowmaker, "Tetrachromatic Color Vision in the Duck: Microspectrophotometry of Visual Pigments and Oil Droplets", *Journal of Comparative Physiology A*. T. 162, 1988, páginas 225-235; A. Palacios, S. Bloch, C. Martinoya, F. Varela, "Color Mixing in the Pigeon", *Vision Research*, t. 30, 1990, P. 587-596.

[13] F. Harosi, Y. Hasimoto, "Ultraviolet Visual Pigment in a Vertebrate: A Tetrachromatic Cone System in the Dace", Science, t. 222, 1983, p. 1021-1023; C. Neumeyer, *Das Farbensehen des Goldfisches*, Hab. Thesis, Universidad Mainz, 1986.

[14] Para el caso de los peces dorados y la tortuga ver, F. Varela, A. Palacios, "Tetra-chromacy: An Analysis of Color Hyperspaces", *Biological Cybernetics*, presentado.

nuestra experiencia cromática dentro de dicho dominio sin un remanente. Las diferencias entre las dimensiones cromáticas son de la naturaleza de la *inconmensurabilidad.*

Ya he dicho que el espacio del color define lo que el color es. Entonces, la consecuencia de la línea comparativa de análisis es que, el color, como fenómeno, es inseparable de una variedad de encarnaciones. ¿Debemos entonces caer en una total relatividad acerca del color y tratarlo como una *qualia* subjetivista?[15] Habiendo procurado evitar hasta ahora el *Scila* del objetivismo, al parecer caemos en el *Caribdis* del subjetivismo. Para abordar este otro extremo debemos proceder a la segunda parte de nuestro argumento comparativo, el cual puede expresarse de la siguiente manera:

(c) La estructura de un receptor de color dado implica cierto camino evolutivo y nicho ecológico y es en referencia a ambos que se encarna el color.

(d) La posición subjetivista es una que se abstrae a sí misma del incrustamiento mutuo de los organismos y su entorno. Para comprender plenamente al color, debemos comprender las muchas variaciones de los distintos tipos de dominios cromáticos. Ilustremos este punto con dos ejemplos. El primero es la observación reciente de que entre los simios del nuevo mundo todos los machos son dicromatos y tres cuartos de las hembras son tricromatos.[16] Pueden invocarse varias hipótesis para explicar este polimorfismo: selección grupal (diversificación del énfasis perceptual), equilibrio ecológico (diversificación de las fuentes de alimentos). El segundo ejemplo es de la bien conocida visión de color de las abejas, que evolucionó, respecto a la de los humanos, una forma relativamente diferente de tricromaticidad. En este caso, el dominio de color contiene canales oponentes que se extienden a la gama de los UV sin

[15] Para obtener una argumentación reciente en este sentido, ver C. L. Hardin, *Color for Philosophers*, Cambridge, Hacket Publishing, 1988.

[16] G. Jacobs, J. Neitz, M. Crognale, "Color vision Polymorphism and its Photopigment Basis in a Callitrichid Monkey", *Vision Research*, t 27, 1987, page 2089-2100.

sensibilidad en las longitudes de onda más largas, es decir, se trata de un dominio tricromático desplazado[17].

El punto es que el subjetivismo no puede explicar estos tipos de fenómenos porque se centra en el sujeto perceptor individual, y por ende ignora el nivel ecológico apropiado de la explicación. Los dos casos mencionados requieren que veamos al sustrato neurobiológico y los canales de color relevantes en una especificación recíproca de lo que cuenta como un nicho para el organismo. es necesario retener que este argumento se aplica con igual fuerza a la posición de Locke, en el sentido de que el color es la reflectancia de los objetos, ya que se refiere al mundo desnudo de la física y no a la distinción ecológicamente inserta, relevante que emerge en una historia evolutiva.

De este modo llegamos a la conclusión de nuestro argumento comparativo: el color es siempre relativo a la estructura e historia de un receptor encarnado ecológicamente. Para explicar el color, debemos *generalizar sobre la base de perceptores ecológicamente insertos*. Ahora discutiremos en mayor detalle lo que queremos decir con el término "ecológicamente inserto".

3. Una visión enactiva del color

Resulta claro que lo que hemos llamado aquí el punto de vista objetivista del color calza bien con la sensibilidad del entendimiento y la percepción como una forma de procesamiento de la información: la luz golpeando la retina y dando lugar a señales que surgen de un estado dado de las cosas en el mundo. Una debilidad de esta posición es que, aun si pone un gran énfasis en el procesamiento interno, el punto de referencia fundamental es un medio ambiente dado. Esto descuida completamente la observación fundamental de que el medio ambiente para un sistema viviente no es pre-dado sino especificado y formado junto con su historia evolutiva, La visión

---

[17] R. Menzel, "Spectral Sensitivity and Color Vision in Invertebrates" en H. Autrum, ed. *Handbook of Sensory Physiology*, vol. VII/6ª, Berlin, Springer, 1979.

no es una recuperación de características dadas previamente, sino el ordenamiento de un mundo posible. Un mundo visual no es ni encontrado ni inventado, sino *ordenado*. Huelga decir que una explicación más acabada de este punto excede los límites de este trabajo, pero me referiré a él en relación con el tema de la visión del color[18].

La tentación desde nuestra inercia de ver la percepción como la representación de un mundo pre-dado equivale a decir: "Quizás deberíamos concluir, como ustedes argumentan, que el color no puede ser identificado con la reflectancia de la superficie, pero la visión del color es aún una forma más o menos perfeccionada de recuperar la reflectancia de la superficie". Existen varios problemas con este punto de vista. En primer lugar, como mencionamos antes, el color aparece en varios tipos de nichos. Los peces de aguas profundas tienden a ser dicromatos, en tanto que aquellos que viven cerca de la superficie tienden a ser tricromatos y tetracromatos[19]. ¿Cuál es la norma para juzgar aquí cuál es la reflectancia adecuada por recuperar? Uno podría decir nuevamente que esto sólo muestra que diferentes nichos representan a diferentes regularidades en el mundo externo que son optimizadas, como pensaba Marr[20]. La debilidad de esta posición (y es una gran debilidad) es que esa dependencia en los modelos de optimicidad de la relación organismo-medio ambiente nos conducen a la espesura de los debates acerca de la validez de la optimicidad como un argumento evolucionista[21].

Sin embargo, nos gustaría volver a un punto mencionado al final de la primera Sección. La reflectancia de la superficie no es

---

[18] Para mayores detalles de esta argumentación, ver F. Varela, *Connaître: les sciences cognitives*, Paris, Seuil, 1988; H. Maturana, F. Varela, *The Tree of Knowledge*, Boston, New Science Library, 1987.

[19] E. Macnichol, *The Ecology of Vision*, Nueva York, Oxford University Press, 1982.

[20] D. Marr, *Vision*, San Francisco, Freeman, 1982. Ver también P. Kitcher, "Marr's Computational Theory of Vision", *Philosophy of Science*, T. 55, 1988, páginas 1-24.

[21] Para mayor información a este respecto, ver J. Dupré, ed. *The Latest on the Best*, Cambridge, MIT Press, 1987.

algo que esté dado en la naturaleza. ¿Cómo ha de especificar uno lo que es un borde, un límite y orientación, excepto en referencia a algún sistema visual para el cual estas distinciones sean relevantes? ¿Cuáles son las áreas y límites relevantes que separan el brillo del lustre? ¿Cuántos canales cromáticos se activarán para especificar un dominio cromático? Como lo indican las investigaciones recientes en A.I., estas son preguntas difíciles y espinudas. No hay solución, salvo a través de una *finalización relativa a un sistema de percepción estándar*. Este observador estándar resulta claro para el ingeniero A.I. pero no para el biólogo comparativista, para quien los alcances de la naturaleza son más amplios. El punto no es que las superficies/color/dimensiones sean subjetivas, sino que los objetos/colores/bordes suponen simultáneamente la existencia de un perceptor para quien algo cuenta como una sola cosa, y estos vienen en distintas variedades.

Postulamos que en la historia, las redes entrelazadas de las especies han ido danzando en un juego evolutivo.Tal danza, en la cual la superficie y el color han desempeñado un papel central, ha producido, además, varias maneras de colorear y no resulta posible definir un estándar universal. Las maneras de colorear se definen a sí mismas como las parejas que bailan. De hecho, la historia natural de esta estética con valor policentrista está aún por escribirse.

De este modo llegamos a la conclusión de este artículo, en el que hemos tratado de mostrar de que manera el color y su dimensión comparativa en particular pueden verse como un caso de estudio en las bases de la ciencia cognitiva. Nuestra conclusión es que proporciona evidencia para lo que hemos llamado una visión enactiva de la cognición que evita dos extremos:

- En primer lugar, el extremo de la cognición como referida a (representando) un mundo que es pre-dado, incluso considerando dicho procesamiento interno y estado incompleto.

- En segundo lugar, el extremo de la cognición entendida como una variante del constructivismo que desconoce la incrustación ecológica del organismo y sus coacciones.

Nuestra posición es que la cognición es *enacción,* un mutuo pliegue/despliegue natural del organismo y el mundo revelado a través de las regularidades que se dan a la luz, tal como el color. Las diversas maneras de colorear son un excelente ejemplo: nosotros, animales de esta tierra, vivimos en nuestros diversos espacios de color y por ende nuestros cerebros y nuestros universos se encuentran. A Cézanne probablemente le hubiera encantado esto.

¿DE DÓNDE VIENE EL SIGNIFICADO PERCEPTUAL?
Cartografía de conceptos actuales

Este texto tiene el mérito de presentar mi visión enactiva del conocer de manera progresiva y cuasi-histórica a partir de los orígenes de las ciencias cognitivas. La diferencia entre cognitivismo o el paradigma inicial, conexionismo como sucesor triunfante, y la extensión a la visión enactiva, es crucial para comprender los nudos filosóficos en juego. Aunque escrito hace ya diez años, nada de lo que ha pasado entretanto ha puesto en cuestión lo que se esboza aquí y, lo que es más, el paradigma "encarnado" o enactivo no ha dejado de avanzar y conquistar más adeptos.

1. Introducción

Aclaraciones

Este ensayo fue escrito con el propósito de proporcionar una base mínima común de discusión. Necesariamente es un intento ambicioso de dar cuenta de manera concisa de las diversas ideas actualmente vigentes acerca del origen del significado en los sistemas vivos y artificiales, de modo que sea accesible a un público interdisciplinario, pero a la vez lo suficientemente sustantivo como para generar un debate entre los especialistas. Desde ya presento mis disculpas a ambos grupos por pasajes que puedan parecer irritantemente simples o demasiado oscuros.

Asimismo, me he limitado a las habilidades cognitivas básicas o "de bajo nivel", esto es, a temas más cercanos a la percepción,

---

(Varela,F. (1991), Whence the origin of perception? A cartography of current ideas, in: F.Varela and J.P.Dupuy (Eds.) *Understanding Origin: Contemporary Ideas on the Origin of Life, Mind and Society*, Boston Studies Phil.Sci., Kluwer, Boston, pp.235-265).

movimiento y el simple aprendizaje. Esto contrasta con las habilidades cognitivas "de alto nivel", que son temas más cercanos al lenguaje y el razonamiento.

Esta presentación no puede ser neutral y mis preferencias se apreciarán de manera explícita en el texto. Específicamente, argumentaré que la llave maestra de lo cognitivo es su capacidad de *generar* significados: la información no es preestablecida como un orden dado, sino que surgen regularidades a partir de una co-determinación de las actividades cognitivas mismas.

Reseña

La ciencia cognitiva tiene poco más de cuarenta años de existencia. *No* está establecida como una ciencia madura con una dirección clara y un gran número de investigadores que conformen una comunidad, como es el caso de los físicos atómicos o la biología molecular. Por ende, el desarrollo futuro de la ciencia cogniscitiva dista mucho de estar claro, aunque lo que ya se ha producido ha tenido un profundo impacto y esto continuará siendo así. Pero el progreso en este campo se basa en apuestas conceptuales audaces (algo parecido a tratar de colocar al hombre en la luna... sin saber dónde está la luna). Para ser concreto, el conjunto de disciplinas que menciono aquí como parte integrante de la ciencia cognitiva actual son: la neurociencia, la Inteligencia Artificial, la psicología cognitiva, la lingüística y la epistemología.

El principal objetivo de este escrito de información básica es proporcionar una imagen radiográfica del estado actual de la ciencia cognitiva en relación con la percepción y el origen del significado. Ahora bien, al igual que cualquiera que haya examinado alguna vez una disciplina científica con detención, he visto que las ciencias cognitivas constituyen una *diversidad* de visiones semicompatibles y no un campo monolítico. Más aún, como cualquier actividad social, tiene polos de dominación, de tal manera que algunas de las voces participantes adquieren más fuerza que otras en distintos períodos de tiempo. Esto es especialmente cierto en la revolución

de la ciencia cognitiva moderna, que fue fuertemente influenciada por algunas líneas de investigación, particularmente en los Estados Unidos. Mis preferencias en este documento apuntan al énfasis de la diversidad.

Procederé en cuatro etapas conceptual y prácticamente distintas entre sí.

- Etapa 1: Una mirada a los años fundacionales;
- Etapa 2: Símbolos: el paradigma cognitivista;
- Etapa 3: Emergencia: alternativas a la manipulación de los signos;
- Etapa 4: Enacción: alternativas a las representaciones.

A través de esta descripción cuadricular y sus articulaciones, examinaremos las bases de lo que ya ha sido establecido como una clara tendencia (Etapas 1 y 2) y veremos que este paradigma establecido coexiste con un espectro más amplio de puntos de vista (Etapas 3 y 4). Esta provocadora heterodoxia tiene el potencial de generar profundos cambios.

2. Una mirada a los años fundacionales

Comenzaremos con una breve mirada a las raíces de estas ideas en la década de 1943-1953 de tal manera que podamos tocar temas que sean relevantes para nosotros hoy en día[1]. De hecho, prácticamente todos los temas que se debaten activamente en la actualidad ya fueron introducidos en estos años formativos, lo cual demuestra que son temas profundos y difíciles de abordar. Los 'fundadores' sabían muy bien que sus preocupaciones implicaban una nueva ciencia y la bautizaron con un nuevo nombre: *cibernética.* Este término no se usa en la actualidad y muchos científicos

---

[1] Esta sección debe mucho a nuestro reciente trabajo colectivo acerca de la historia olvidada de la cibernética temprana, la auto-organización, el entendimiento, publicados como Cahiers du CREA Números 7 y 9. La única otra fuente útil es S. Heims, *John von Neumann and Norbert Wiener,* MIT Press, 1980. El reciente libro de H. Gardner, *The Mind's New Science: A History of The Cognitive Revolution,* Basic Books, 1985, discute este período solamente de manera superficial.

cognitivos de hoy ni siquiera estarían dispuestos a admitir un parentesco con él. Referirse a esto resulta útil, ya que nos indica que para establecerse como una ciencia, en su orientación claramente cognitivista (Etapa 2 de este texto), la futura ciencia cogniitiva debía cortar con sus raíces, las cuales son más complejas y borrosas, pero también más ricas. Esto muchas veces es así en la historia de la ciencia: es el precio que debe pagarse por el paso desde una etapa exploratoria a la de un programa de investigación, digamos desde lo borroso a lo cristalino.

Los frutos del movimiento cibernético

La fase cibernética de la ciencia cognitiva produjo una impresionante gama de resultados concretos, aparte de su prolongada influencia a través del tiempo (muchas veces subterránea). Algunos de estos son:

- uso de la lógica matemática para comprender el funcionamiento del sistema nervioso;
- invención de máquinas procesadoras de datos (tales como los computadores), sentando así las bases para la Inteligencia Artificial.
- establecimiento de la metadisciplina de la teoría de los sistemas, la cual ha dejado su huella en muchas ramas de la ciencia, tales como la ingeniería (análisis de sistemas, teoría del control), biología (fisiología reguladora, ecología), ciencias sociales (terapia familiar, antropología estructural, estudios urbanos) y economía (teoría de los juegos);
- la teoría de la información como una teoría estadística acerca de las señales y canales de comunicación;
- los primeros ejemplos de los sistemas autoorganizativos.

La lista es impresionante. Tendemos a considerar muchas de estas nociones y herramientas como una parte integral de nuestras vidas, sin embargo no existían antes de esta década formativa y

todas fueron el resultado de un intenso intercambio entre personas de muy distintos ámbitos: un esfuerzo multidisciplinario con un éxito extraordinario.

La lógica y la ciencia de la mente

La intención declarada del movimiento cibernético era crear una *ciencia de la mente.* Para sus seguidores, los fenómenos de la mente habían estado demasiado tiempo en manos de psicólogos y filósofos y se sintieron llamados a mencionar los procesos que subyacen tras los fenómenos mentales en la forma de mecanismos explícitos y fórmulas matemáticas[2].

Una de las mejores ilustraciones de este modo de pensar fue el artículo 'A logical calculus immanent in nervous activity'[3] (1943), escrito por McCulloch y Pitts, el cual originó un sinúmero de trabajos sobre el tema. En este artículo se dieron varios grandes pasos. En primer lugar, se propuso la *lógica* como la disciplina adecuada para comprender al cerebro y la actividad mental. En segundo lugar, se vio el cerebro como el dispositivo que *encarna* principios lógicos en sus elementos componentes o neuronas. Cada neurona era vista como un dispositivo tipo umbral que podía ser activo o inactivo. Estas sencillas neuronas podían entonces conectarse unas a otras y mediante sus interconexiones llevar a cabo las funciones de operaciones lógicas de tal manera que el cerebro entero puede considerarse como una máquina deductiva.

Estos planteamientos fueron muy importantes para la invención de los computadores digitales[4]. En esa época se usaban

---

[2] Las mejores fuentes en este caso son las muchas veces citadas Macy-Conferences, publicadas como *Cibernetics-Circular causal and feedback Mechanisms in Biological and Social Systems,* Josiah Macy Jr. Foundation, Nueva York, 5 volúmenes.

[3] *Bulletin of Mathematical Biophysics,* 5, 1943. Reimpreso en W. McCulloch, *Embodiments of Mind,* MIT Press, 1965.

[4] Para un interesante punto de vista acerca de este momento histórico/conceptual, ver también A. Hodges, *Alan Turing: The Enigma of Intelligence,* Touchstone, Nueva York, 1984.

tubos al vacío para implementar las neuronas de McCulloch-Pitts, mientras que hoy se utilizan chips, pero las computadoras modernas se construyen sobre la base de la misma arquitectura de von Neumann. Este gran salto tecnológico también sentó las bases para el enfoque predominante en el estudio científico de la mente, cuya culminación sería, en la siguiente década, el paradigma cognitivista.

El fin de una era

Por cierto que hubo mucho más en esta década creativa. Por ejemplo, estaba el debate sobre si la lógica era de hecho suficiente para comprender el cerebro, ya que ésta descuidaba sus cualidades distributivas. Se plantearon teorías y modelos alternativos, los cuales en su mayoría habrían de permanecer latentes hasta que fueran revividos y pasaran a constituir una importante alternativa para la ciencia cognitiva en los años 70 (Etapa 3). En 1953, contrastando con su vitalidad y unidad iniciales, los principales actores de la fase cibernética se habían ido distanciado unos de otros y muchos murieron poco tiempo después. No obstante, la idea de la mente como un cálculo lógico habría de seguir desarrollándose.

3. Símbolos: Las hipótesis cognitivas

Así como 1943 fue claramente el año en que nació la fase cibernética, 1956 fue el año que vio nacer la segunda fase de la CC. Durante ese año, en dos reuniones celebradas en Cambridge y Dartmouth respectivamente, surgieron nuevas voces (como las de Herbert Simon, Noam Chomsky, Marvin Minsky y John McCarthy) que propusieron las ideas que se convertirían en las principales pautas de la ciencia cognitiva moderna[5].

La idea central es que la inteligencia (incluyendo la inteligencia humana) se asemeja de tal manera a la computadora en sus

[5] Ver H. Gardner, *op.cit.*, Capítulo 5 para este período.

características esenciales que la cognición puede *definirse* como las computaciones de representaciones simbólicas. Es claro que esta orientación no hubiera podido surgir sin las bases establecidas en la década anterior y que han sido evocadas en la sección dos. La principal diferencia es que una de las muchas ideas tentativas originales se propone aquí como una *hipótesis* completa, con un fuerte deseo de fijar sus límites más allá de sus raíces, que son de carácter más amplio, exploratorias e interdisciplinarias, y donde las ciencias biológicas y sociales figuraban de manera preeminente, en toda su variada complejidad. El *cognitivismo*[6] es un nombre que resulta conveniente para identificar esta orientación amplia aunque bien delineada, y que ha motivado numerosos desarrollos científicos y tecnológicos desde 1956 en todas las áreas de la ciencia cognitiva.

Resumen de la doctrina

El programa de investigación cognitiva puede resumirse como las respuestas a las siguientes preguntas:

Pregunta 1: ¿Qué es la cognición?

Respuesta: Procesamiento de información: manipulación de símbolos sobre la base de normas.

Pregunta 2: ¿Cómo funciona?

Respuesta: A través de cualquier dispositivo que pueda sustentar y manipular elementos físicos discretos: los símbolos. El sistema interactúa solamente con la forma de los símbolos (sus atributos físicos), no con su significado.

Pregunta 3: ¿Cómo sé cuando un sistema cognitivo está funcionando adecuadamente?

---

6 Esta designación se justifica en J. Haugland (editor), *Mind Design*, MIT Press, 1981. Otras designaciones utilizadas son: computacionalismo (Fodor) o procesamiento simbólico. Para esta sección me he beneficiado mucho del artículo de D. Andler aparecido en *Cahiers du CREA N°9*.

Respuesta: Cuando los símbolos representan apropiadamente algún aspecto del mundo real y el procesamiento de la información conduce a resolver con éxito el problema que enfrenta el sistema.

Obviamente, el programa cognitivista, tal como se lo resume arriba, no surgió como algo ya hecho, como Atena del cerebro de Zeus. Lo presentamos aquí con el beneficio de 30 años de visión retrospectiva. Sin embargo, este audaz programa de investigación no solamente se ha establecido plenamente, sino que incluso hoy en día es *identificado* por muchos con la ciencia cognitiva misma, aunque esto está cambiando rápidamente. Hasta hace muy poco, sólo unos cuantos de entre sus participantes activos y nadie del público en general, eran sensibles a sus orígenes o sus actuales desafíos y alternativas. "El cerebro procesa información del mundo exterior" es una frase cotidiana que todo el mundo entiende. Resulta extraño tratar afirmaciones de este tipo como problemáticas, en lugar de hacerlo como algo evidente, y por ello la conversación que siga será inmediatamente catalogada como 'filosófica'. Esta es una forma de *ceguera* introducida en el sentido común con posterioridad al establecimiento del cognitivismo.

Lo que ha forjado el cognitivismo: Inteligencia Artificial

Las manifestaciones del cognitivismo nunca son más visibles que en la Inteligencia Artificial, que es la *interpretación lógica* de la hipótesis cognitivista. Con el correr de los años, se han logrado muchos avances teóricos y se han desarrollado aplicaciones tecnológicas interesantes dentro de esta orientación: sistemas expertos, robótica, procesamiento de imágenes. Estos resultados han sido publicitados ampliamente y no necesitamos insistir con ejemplos ahora.

Psicología cognitiva

La hipótesis cognitiva encuentra su interpretación más literal en la Inteligencia Artificial. Su labor complementaria es el estudio de los

sistemas cognitivos naturales, biológicamente implementados, muy especialmente el hombre. También en este caso, las representaciones susceptibles de ser caracterizadas computacionalmente han sido la herramienta explicativa principal. Las imágenes mentales se consideran como eventos en un sistema formal y es la actividad de la mente lo que da a estas representaciones su gama de actitudes: creencias, deseos, planes y así sucesivamente. Entonces, a diferencia de lo que ocurre con la Inteligencia Artificial, en los sistemas cognitivos *naturales*, surge el interés y se asume que sus representaciones cognitivas son intencionales, es decir, son *sobre* algo *para* el sistema[7].

Un buen ejemplo de esta línea de investigación es el siguiente experimento. Se entregaron figuras geométricas a ciertos sujetos, pidiéndoles que las hicieran rotar en su mente. De manera consistente, éstos informaron que la dificultad de la tarea dependía de los distintos grados de libertad con los cuales debía rotarse la figura. Es decir que todo ocurre tal como si tuviéramos un espacio donde las figuras se hacen girar como en una pantalla de televisión[8]. Con el tiempo, estos experimentos produjeron una teoría explícita que postulaba normas en base a las cuales opera el espacio mental, que serían similares a las utilizadas en las pantallas computacionales que operan con información almacenada. Estos investigadores postularon que existe una interacción entre las operaciones que se asemejan a lenguajes y las operaciones que se asemejan a imágenes y que juntas generan nuestro ojo interior[9]. Este enfoque ha generado abundante literatura, tanto a favor como en contra[10], y cada nivel de observaciones ha recibido interpretaciones alternativas. Sin embargo, el estudio del imaginario es un ejemplo perfecto de la forma en que el enfoque cognitivista procede cuando estudia un fenómeno de la mente.

---

7 Más sobre este tema en J. Searle *Intentionality*, Cambridge U.P. 1983.

8 R. Shepard y J. Metzler, *Ciencia* 171, 701-703 (1971).

9 S. Kosslyn, *Psychol. Rev.* **88** 46-66, 1981

10 Ver *Beh, Brain Sci.* **2**, 535-581 (1979)

Procesamiento de información en el cerebro

Otro efecto igualmente importante del cognitivismo es cómo ha dado forma a los puntos de vista actuales acerca del cerebro. Con el correr de los años, casi toda la neurobiología (y su enorme *corpus* de evidencia empírica) se ha visto impregnada del punto de vista del procesamiento de información. La mayor parte de las veces, los orígenes y suposiciones de este punto de vista ni siquiera se cuestionan[11].

El mejor ejemplo de este enfoque está dado por los conocidos estudios sobre el cortex de la vista, en donde uno puede detectar respuestas eléctricas de las neuronas cuando al animal se le muestra una imagen visual. Pronto se informó que era posible clasificar a estas neuronas corticales como 'detectores de características' que responden a ciertos atributos del objeto presentado: orientación, contraste, velocidad, color y así sucesivamente[12]. En la misma línea de la hipótesis cognitivista, estos resultados fueron vistos como una entrega de sustancia biológica a la idea de que el cerebro recoge información visual de la retina a través de neuronas de funcionalidad específica que hay en el cortex. Luego, esta información es traspasada a etapas posteriores en el cerebro para un mayor procesamiento (categorización conceptual, asociaciones de memoria y eventualmente, acción).

En su forma más extrema, este concepto del cerebro se manifiesta en la teoría de la célula abuela de Barlow,[13] en la que existe una correspondencia entre conceptos o preceptos y las neuronas. (Es el equivalente, en Inteligencia Artificial, a los detectores y las líneas etiquetadas).

---

[11] Esta es la frase inicial de un libro de texto popular en la neurociencia: "El cerebro es un conjunto de células sin descanso que constantemente recibe información, la elabora y la percibe, y toma una decisión". S. Kuffler y J. Nichols, *From Neuron to Brain*, Sinauer Associates, Boston, Segunda Edición, 1984, página 3.

[12] D. Hubel y T. Wiesel, J. Physiology **160**, 106 (1962) Para un análisis reciente de este trabajo, ver Kuffler y Nichols, *op.cit.* Ch 2-4.

[13] H. Barlow, *Perception*, 1 371-392.

Breve resumen de las posturas disidentes

La ciencia cognitiva-como-cognitivismo constituye un programa de investigación bien definido, el cual incluye instituciones prestigiosas, publicaciones periódicas, así como instituciones relacionadas con la tecnología aplicada y entidades comerciales internacionales. La mayor parte de las personas que trabaja dentro de la ciencia cognitiva adheriría —consciente o inconscientemente— al cognitivismo o a sus variaciones más próximas. Después de todo, si los principales ingresos de una persona provienen de la escritura de programas para la representación del conocimiento o la búsqueda de neuronas para tareas bien definidas ¿cómo podría ser de otro modo? Para efectos de lo que nos interesa aquí, es importante llamar la atención sobre la profundidad de este *compromiso social* que afecta a un gran sector de la comunidad investigadora de la ciencia cognitiva. Ahora nos concentraremos en la posición disidente, que adopta básicamente dos formas:

- Una crítica a las computaciones simbólicas en tanto portadoras apropiadas de las representaciones.
- Una crítica a la idoneidad de la idea de las representaciones como el punto de Arquímedes de la ciencia cognitiva.

4. Emergencia: Alternativas a los símbolos

Orígenes de las ideas de autoorganización

Durante la década formativa, ya se habían propuesto y discutido alternativas a la preponderancia del dominio de la lógica como principal enfoque para la ciencia cognitiva. En las conferencias Macy, por ejemplo, se alegó que en los cerebros mismos no existen normas ni un procesador de lógica central ni su información se encuentra almacenada en direcciones precisas. En lugar de eso, los cerebros parecen operar sobre la base de interconexiones masivas, de una manera distribuida, de modo que su misma conectividad cambie como resultado de la experiencia. En síntesis, presentan una capacidad de autoorganización que no se encuentra en la lógica.

En 1958, F. Rosenblatt construyó el 'Perceptron', un dispositivo simple con cierta capacidad de reconocimiento, basándose únicamente en los cambios de conectividad que se producen entre componentes parecidos a las neuronas[14]. Asimismo, W.R. Ashby realizó el primer estudio de la dinámica de sistemas de gran tamaño con interconexiones al azar, demostrando que exhiben comportamientos globales coherentes[15].

Históricamente, estos puntos de vista alternativos fueron marginados de la escena intelectual en favor de las ideas computacionales discutidas arriba. Fue solamente a fines de los setenta que hubo un explosivo renacer de estas ideas, luego de 30 años de preeminencia de la ortodoxia cognitivista; lo que D. Dennet[16] ha llamado computacionismo de raíz ortodoxa. Ciertamente uno de los factores que contribuyen a este interés renovado fue el redescubrimiento paralelo de ideas autoorganizativas en la física y las matemáticas no lineales.

Motivación para buscar una alternativa

La motivación para dar una segunda mirada a la autoorganización se basó en dos deficiencias del cognitivismo, ampliamente reconocidas. La primera es que el procesamiento de información simbólica se basa en normas secuenciales, aplicadas una a la vez. Este famoso cuello de botella de von Neumann es una limitación dramática cuando la tarea en cuestión requiere grandes cantidades de operaciones secuenciales (tales como el análisis de imágenes naturales o la predicción del clima). Una búsqueda permanente de algoritmos de procesamiento *paralelo* en las arquitecturas clásicas no ha tenido mucho éxito debido a que la filosofía computacional entera evoluciona precisamente en sentido contrario.

---

14 F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and The Theory of Brain Mechanisms*, Spartan Book, 1962.

15 Para mayores informaciones acerca de los complejos orígenes iniciales de las ideas de la auto-organización ver I. Stengers, *Cahier du CREA* N°8, páginas 7-105.

16 "La geografía lógica de los enfoques computacionales", Conferencia MIT Sloan, 1984.

La segunda limitación importante es que el procesamiento simbólico es *localizado:* la pérdida de cualquier parte de los símbolos o normas del sistema implica una falla importante en el funcionamiento. Por el contrario, una operación *distribuida* es altamente deseable, de manera que haya al menos una equipotencialidad relativa y una inmunidad a las mutilaciones.

Estas dos desviaciones del cogcitivismo pueden expresarse de la misma manera: las arquitecturas y los mecanismos están a una gran distancia de la biología. Las tareas visuales más ordinarias, realizadas aún por pequeños insectos, se realizan más rápidamente de lo que es físicamente posible cuando se las simula de una manera secuencial, la elasticidad de la mente frente al daño sin comprometer la totalidad de su competencia ha sido conocida por los neurobiólogos desde hace mucho tiempo.

¿Qué es la emergencia?

Lo anterior sugiere que en lugar de enfocarse en los símbolos como punto de partida, uno podría comenzar con componentes simples (no cognitivos) que se conectarían unos con otros de maneras densas. De acuerdo con este enfoque, cada componente opera solamente en su entorno *local*, pero debido a la estructura de red del sistema completo, existe una cooperación global que *emerge* espontáneamente cuando los estados de todos los componentes participantes alcanzan un estado mutuamente satisfactorio, sin necesidad de una unidad de procesamiento central que guíe la operación completa[17]. Este paso desde normas locales a una coherencia global es el corazón de lo que solía llamarse autoorganización durante los años fundacionales[18]. Hoy en día, se prefiere hablar ya sea de propiedades emergentes o globales, de dinámicas de redes o incluso de sinergia. Si bien no existe una teoría formal unificada acerca de las propiedades emergentes, la teoría regional más evidente

---

17 Para una extensa discusión sobre este punto, ver P. Dumouchet y J-P Dupuy (Editores) *L'Auto-organización: De la physique au politique*, Eds. du Seuil, Paris, 1983.

18 Ver por ejemplo, H. von Foerster (Ed.) *Principles of Self-Organization*, Pergamon Press, 1962.

es aquella de los atractores en la teoría de los sistemas dinámicos[19]. Estos no son propiedad de componentes individuales, sino del sistema completo y sin embargo, cada componente contribuye a su emergencia y características.

Un cambio de perspectiva respecto al cerebro

Investigaciones recientes han producido evidencia detallada acerca de como las propiedades emergentes están en el núcleo del funcionamiento del cerebro. Esto no resulta tan sorprendente si uno examina los detalles de la anatomía del cerebro. Por ejemplo, si bien es cierto que las neuronas en el córtex visual responden claramente a 'características' específicas de los estímulos visuales, como se mencionó arriba, esto es válido en un animal anestesiado en un ambiente altamente simplificado (interno y externo). Cuando se permite un entorno sensorial normal donde el animal es estudiado mientras se encuentra despierto, en su comportamiento habitual, se ha visto que las respuestas neuronales estereotipadas descritas antes se vuelven altamente sensibles al contexto. Por ejemplo, existen distintos efectos derivados de la inclinación del cuerpo [20] o estímulo auditivo[21]. Más aún, las características de las respuestas neuronales dependen directamente de neuronas localizadas a gran distancia de sus campos receptivos[22].

Así, se ha vuelto cada vez más necesario estudiar las neuronas como miembros de *grandes conjuntos* que constantemente desaparecen y surgen a través de interacciones cooperativas en donde cada neurona tiene múltiples y cambiantes grados de respuesta a los estímulos visuales, dependiendo del contexto. Aun en el extremo

---

[19] Una introducción accesible a la teoría moderna de los sistemas dinámicos es: R. Abraham y C. Shaw, *Dynamics: The Geometry of Behavior*, Aerial Press, Santa Cruz, 3 Volúmenes, 1985.

[20] G. Horn y R. Hill, *Nature* 221, 185-187(1974).

[21] M. Fishman y C. Michael, *Vision Res.*, **13**, 1415 (1973) y F. Morell, ***nature***, **238**, 44-46 (1972).

[22] J. Allman, F. Miezen y E. McGuiness, *Ann. Rev. Neuroscien.* **8** 407-430 (1985)

más periférico del sistema visual, las influencias que el cerebro recibe del ojo se encuentran con una mayor actividad que *desciende* del córtex. Es como resultado del encuentro de estos dos conjuntos de actividad neuronal que emerge una nueva y coherente configuración, dependiendo de la coincidencia o falta de coincidencia entre la actividad sensorial y el ambiente 'interno' en el cortex[23]. En general, una neurona individual participa en muchos de estos patrones globales y muestra una muy escasa significación cuando se la considera individualmente.

Aun cuando estos ejemplos se toman del dominio de la visión para enfartizar el contraste con el ejemplo de la sección anterior, existen muchos otros análisis detallados[24]. No es necesario que insistamos más sobre este punto.

La estrategia (neo)conexionista

Bajo esta orientación alternativa, el cerebro, una vez más, ha sido fuente principal de metáforas e ideas para otros campos de la ciencia cognitivista. En lugar de partir de descripciones simbólicas abstractas, uno parte por toda una gama de componentes absurdos, los cuales, apropiadamente interconectados, pueden tener interesantes propiedades a nivel global. Estas propiedades globales son las que encarnan/expresan las capacidades cognitivas que se buscan.

Por tanto, todo este enfoque depende de la introducción de las conexiones apropiadas y esto por lo general se hace a través de una norma para el *cambio gradual* de las conexiones, partiendo de un estado inicial bastante arbitrario. En la actualidad disponemos de varias de estas normas, pero la más explorada es la Norma de Hebb, según la cual los cambios de conectividad en el cerebro podrían surgir a partir del grado de actividad *coordinada* entre las neuronas: si dos

---

[23] F. Varela y W. Singer, *Exp. Brain Res.* 66, 10-20 (1987).

[24] Una interesante colección de ejemplos es: G. Palm y A. Aersten (Editores) *Brain Theory*, Springer Verlag, 1986.

neuronas tienden a ser activas juntas, su conexión se refuerza, de lo contrario se disminuye. Por lo tanto, la conectividad del sistema se vuelve inseparable de *la historia de su transformación*, relacionada a su vez con el tipo de tarea definida para el sistema. Dado que la acción real ocurre al nivel de las conexiones, se ha propuesto el nombre de (neo)*conexionismo* para esta corriente de investigación[25].

Uno de los factores importantes para el explosivo interés que ha habido en este enfoque fue la introducción de algunos métodos efectivos para hacer un seguimiento de los cambios en las redes, y sobre todo las medidas estadísticas que dotan al sistema de una función de 'energía' global que asegura su convergencia[26]. Por ejemplo, tome N elementos simples, semejantes a neuronas, conéctelos recíprocamente y déles una norma tipo Hebb. Luego dé a este sistema una sucesión de patrones (no-correlativos) en algunos de sus nodos y en cada entrega permita que el sistema se reorganice, reordenando sus conexiones siguiendo su gradiente de energía. Luego de la fase de aprendizaje, cuando se presenta una vez más al sistema uno de estos patrones, lo reconoce, en el sentido de que recae en un atractor único y una configuración interna que se dice representa al ítem aprendido. El reconocimiento es posible siempre y cuando el número de patrones presentado no sea mayor de alrededor de $0.15\ N$. Además, el sistema ejecuta un reconocimiento correcto, incluso si al patrón se le da un ruido adicional o el sistema es parcialmente mutilado[27].

Otra técnica importante preferida por algunos investigadores es la propagación retrógrada: cambios en las conexiones neuronales

---

[25] El nombre se propone en: J. Feldman y D. Ballard 'Connectionist Models and their properties', *Cognitive Science* 6 205-254 (1982). Para una discusión extensa del trabajo actual en esta dirección ver: D. Rumelhart y J. McClelland (Editores) *Parallel Distributed Processing: Studies on the Microstructure of Cognition*, MIT Press, 1986, 2 volúmenes.

[26] La idea principal se debe a J. Hopfield, *Proc. Natl. Acad. Sci.* (USA), 79, 2554-2556 (1982).

[27] Existen muchas variantes asociadas a estas ideas. Ver en particular: G. Hinton, T. Sejnowsky y D. Ackley, *Cognitive Science* 9, 147-163(1984) y G. Toulouse, S. Dehaene y J. Changeaux, *Proc. Natl.. Acad. Sci.* (USA), 83, 1965-1968 (1986).

al interior de la red (unidades ocultas) con el fin de minimizar la diferencia entre la respuesta de la red y lo que se espera de ella, de manera muy parecida a alguien tratando de imitar a un instructor[28]. NetTalk, un famoso ejemplo reciente de este método, es una máquina de conversión grafema-fonema que trabaja frente a algunas páginas de texto en inglés sencillo. Como resultado, NetTalk puede leer en voz alta un nuevo texto en lo que muchos oyentes consideran un inglés deficiente pero comprensible[29].

Los modelos conexionistas proporcionan con impresionante gracia, un modelo operativo para una variedad de capacidades cognitivas básicas. El actual trabajo con esta orientación se justifica por varias razones. En primer lugar, la Inteligencia Artificial cognitivista y la neurociencia tuvieron algunos resultados convincentes en la justificación o reconstrucción de algunos de los comportamientos cognitivos que acabo de describir. En segundo lugar, estos modelos son bastante cercanos a los sistemas biológicos, lo que significa que uno puede trabajar con un grado de integración entre Inteligencia Artificial y neurociencia que hasta ahora era impensable. Finalmente, los modelos son lo suficientemente generales como para aplicarse, con muy pocas modificaciones a una variedad de áreas.

Resumen de la doctrina

Esta orientación alternativa —conexionista, emergente, auto-organizativa, asociativa y basada en la dinámica de redes— es reciente y variada. La mayoría de quienes se identificarían con ella tienen puntos de vista muy distintos acerca de lo que constituye la ciencia cognitiva y también acerca de su futuro. Con esta salvedad en mente, las siguientes son las respuestas de la postura alternativa a las preguntas anteriores:

---

28 La idea se debe a D. Rumelhart, G. Hinton y R. Williams, en: Rumelhart and McClelland, *op cit.*, Cap. 8.

29 T. Sejnowski y C. Rosenbaum, 'Nettalk: A parallel network that learns to read aloud' TR JHU/EEC-86/01, John Hopkins University.

Pregunta 1: ¿Que es cognición?

Respuesta: La emergencia de estados globales en una red de componentes simples.

Pregunta 2: ¿Cómo funciona?

Respuesta: A través de normas locales para el funcionamiento individual y normas para los cambios en la conectividad entre los elementos.

Pregunta 3: ¿Cómo se sabe cuándo un sistema cognitivo está funcionando adecuadamente?

Respuesta: Cuando puede verse que las propiedades emergentes (y la estructura resultante), pueden corresponder a una capacidad cognitiva específica: el éxito en la solución para una tarea requerida.

Abandono de los símbolos

Uno de los aspectos más interesantes de este enfoque alternativo de la ciencia cognitiva es que los símbolos, en su sentido más convencional, no cumplen ninguna función. Esto implica un distanciamiento radical del principio cognitivo básico: la estructura física de los símbolos, su forma, está separada para siempre de lo que representan, su significado. Esta separación entre forma y significado fue el golpe maestro que creó el enfoque computacional, pero también implica una debilidad cuando se abordan los fenómenos cognitivos a un nivel más profundo. ¿Cómo *adquieren* los símbolos su significado? ¿De dónde viene esta actividad adicional que, por interpretación, no está en el sistema cognitivo?

En situaciones en las que el universo de ítemes susceptibles de ser representados es finito y está claramente definido (como por ejemplo cuando una computadora está siendo programada o cuando se está realizando un experimento con un conjunto de estímulos visuales predefinidos), la asignación de significados es clara. Cada ítem físico discreto dentro del sistema cognitivo es asignado a un ítem

externo (su significado referencial), una operación cartográfica que el observador proporciona fácilmente. Si se eliminan estas restricciones, *todo* lo que queda es la forma del símbolo y el significado pasa a ser un fantasma, como si contempláramos patrones de bits en un computador cuyo manual de operaciones se hubiera perdido.

En el enfoque conexionista, el significado está vinculado al desempeño general (ya sea en el reconocimiento o el aprendizaje). Por ende, el significado se relaciona con el estado global del sistema y no se radica en un símbolo o símbolos particulares. La distinción forma/significado a nivel simbólico desaparece, y reaparece con una vestimenta diferente: el observador proporciona la correspondencia entre el estado global del sistema y el mundo al que se supone maneja. Esta es, entonces, una manera radicalmente diferente de trabajar con las representaciones. Volveremos a este tema más adelante.

5. Vinculación de símbolos y emergencia

En esta etapa, la pregunta obvia que se ha de considerar es qué *relación* hay entre los puntos de vista simbólicos y emergentes respecto al origen de propiedades cognitivas simples. La respuesta obvia es que estos dos puntos de vista deberían ser enfoques complementarios arriba-abajo y abajo-arriba o bien que deberían ser pragmáticamente adyacentes en alguna modalidad mixta, o simplemente usarse en diferentes etapas. Un típico ejemplo de esta vía es describir la visión temprana en términos conexionistas, digamos hasta el córtex visual primario, pero asumir que a nivel del córtex inferotemporal, la descripción debería basarse en programas simbólicos. Sin embargo, el status conceptual de dicha síntesis dista mucho de estar claro y aún faltan ejemplos concretos.

En mi opinión, la relación más interesante entre las descripciones emergentes y simbólicas es una de *inclusión,* esto es ver a los símbolos como un nivel superior de descripción de las propiedades insertas en un sistema distribuido subyacente. El caso del llamado código genético es paradigmático y lo usaré aquí para ser más concreto. Durante muchos

años, los biólogos consideraban las secuencias de proteínas como instrucciones codificadas en el ADN. Sin embargo, es claro que las ternas del ADN son capaces de especificar de manera predecible un aminoácido en una proteína si y sólo si están insertas en el metabolismo de la célula, esto es, en medio de miles de regulaciones enzimáticas en una compleja red química. Es solamente por las regularidades emergentes de tal red considerada como un todo que podemos colocar en la misma categoría este entorno metabólico y tratar a las ternas como códigos para aminoácidos. En otras palabras, la descripción simbólica es posible a otro nivel de descripción. Sin duda se puede tratar tales regularidades simbólicas de manera particular, pero su status e interpretación es muy diferente a cuando se las considera conforme a su valor nominal, independientemente del substrato del cual surgen[30].

El ejemplo de la información genética puede transponerse directamente a las redes cognitivas con las cuales tratan los neurocientistas y los conexionistas. De hecho, algunos investigadores han expresado recientemente este punto de vista[31]. En la teoría de la armonía de Smolesnky[32], por ejemplo, los átomos fragmentarios de 'conocimiento' sobre circuitos eléctricos vinculados por algoritmos estadísticos distribuidos, generan un modelo de razonamiento intuitivo en este dominio. La competencia de este sistema completo puede describirse como la realización de inferencias basadas en leyes simbólicas, pero su desempeño se ubica en diferentes niveles y nunca se alcanza por referencia a un intérprete simbólico. Este punto se ilustra gráficamente en la Figura 1.

---

[30] Para obtener una distinción entre la descripción y explicación simbólica y emergente en los sistemas biológicos, ver F. Varela, *Principios de Autonomía Biológica*, North Holland, New York, 1979, Capítulo 7 y más recientemente S. Oyama, *The Ontogeny of Information*, Cambridge University Press, 1985.

[31] Ver D. Hill 'Intelligence as emergent behavior' , *Daedalus*, Winter 1989 y P. Smolensky, 'On the proper treatment of connectionism', Beh. Brain Sci. 11: 1, 1989. En una línea muy diferente, J. Feldman, 'Neural representation of conceptual knowledge' U. Rochester TR189 (1986) propone un punto de vista intermedio entre sistemas de 'puntuación' y sistemas distribuidos.

[32] P. Smolesnky en: Rumelhart and McClelland, *op.cit.*, Capítulo 6.

Entonces, una vinculación productiva entre un cognitivismo menos ortodoxo, que emerja tranquilamente de un procesamiento paralelo distribuido, proporcionado por los enfoques autoorganizacionales, es una posibilidad concreta, especialmente en la Inteligencia Artificial orientada a la ingeniería. Esta posible complementación indudablemente producirá resultados visibles y bien puede convertirse en la tendencia dominante durante muchos años en la ciencia cognitiva.

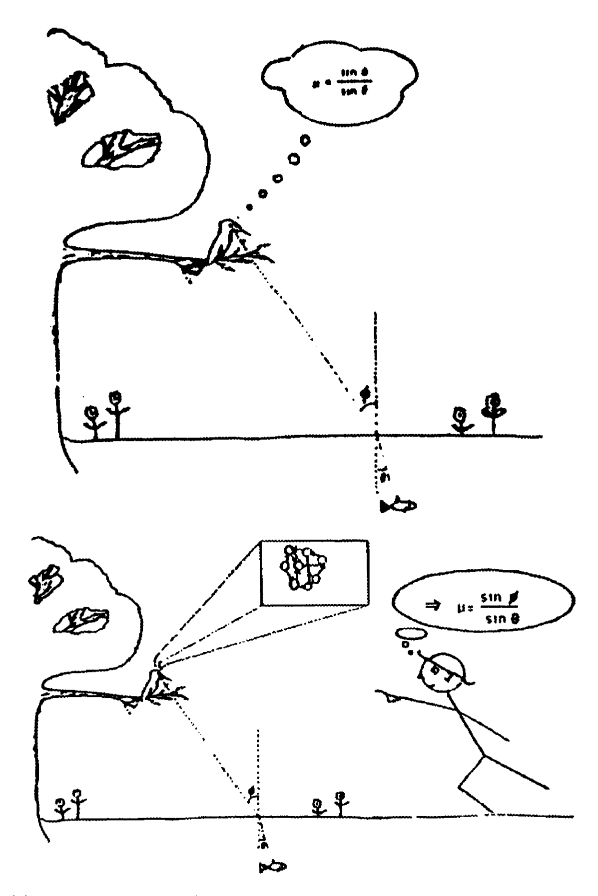


Figura 1. (a) Una caricatura de *Punch* que ilustra de manera suscinta la hipótesis cognitivista. Para atrapar a su presa, este Martín Pescador tiene, en su cerebro, una representación de la ley de refracción de Snell. (b) Otra lectura de la caricatura para indicar cómo los niveles simbólicos pueden verse como surgiendo de la red subyacente.

Por supuesto que este camino es inadmisible desde una postura cognitivista estricta u ortodoxa[33]. Entre los muchos aspectos que cambian desde el punto de vista de la emergencia, dos de ellos merecen ser enfatizados aquí. En primer lugar, la cuestión del origen de un símbolo y su significado (por ejemplo, ¿por qué ATT es un código para alanina?) tiene al menos un enfoque claro para abordarla. En segundo lugar, cualquier nivel simbólico se vuelve altamente dependiente de las propiedades y peculiaridades de la red subyacente y se ve atado a su historia. Por lo tanto, una relación puramente procesal de la cognición, independientemente de sus manifestaciones e historia, es seriamente cuestionada. Estos dos temas nos llevan derecho a nuestra última etapa.

6. ENACCIÓN: ALTERNATIVAS A LAS REPRESENTACIONES

Bases adicionales para la insatisfacción

Resulta tentador detenerse en un análisis de la ciencia cognitiva actual con solamente los dos enfoques antes explicados, pero esto sería inadecuado, ya que en ambas orientaciones (y por ende en alguna posible síntesis futura) *aún* faltarían algunas dimensiones esenciales de la cognición. Es necesario tener en mente un horizonte más amplio para la ciencia cognitiva, nacido de una insatisfacción más profunda que la mera búsqueda de alternativas para los símbolos y también más cercana a las bases mismas de los sistemas representacionistas. Esperemos que este punto de vista, el cual goza en la actualidad de un cierto espacio, no sufra el mismo destino que tuvieron las ideas anteriores sobre la autoorganización, esto es, ser dejado de lado sólo para ser re-descubierto después de treinta años.

---

[33] Esto es discutido exhaustivamente por dos connotados exponentes del cognitivismo: J. Fodor y S. Pylyshin, 'Connectionism and congitive architecture: A critical review', *Cognition*, 1989. Para un análisis de la posición filosófica contraria en *favor* del conexionismo, ver: H. Dreyfus, 'Making a mind vs. modeling the brain: AI again at the cross-roads', *Daedalus*, Winter, 1989.

Insistencia en el sentido común

La total ausencia, hasta ahora, de sentido común en la definición de cognición, constituye la principal fuente de insatisfacción de la llamada alternativa de la Enacción. Tanto en el cognitivismo (por su misma base) como en el conexionismo de nuestros días (por la manera en que se lo practica), todavía se observa que los criterios para la cognición están directamente relacionados con la representación exitosa de un mundo externo, dado previamente, normalmente como una situación que resuelve un problema. Sin embargo, nuestra actividad cognitiva en la vida diaria revela que esta visión de la cognición es demasiado incompleta. Precisamente, la principal habilidad de toda cognición viviente es, dentro de límites amplios, *proponer* los temas relevantes que han de abordarse en cada momento de nuestras vidas. No son pre-dados, sino *activados* o *traídos a un primer plano* desde un contexto y lo que cuenta como relevante es lo que nuestro sentido común determina como tal, siempre de una manera contextual.

Esta es una crítica al uso de la noción de representación como núcleo de la ciencia cognitiva, ya que solamente si existe un mundo pre-dado éste puede ser representado. Si el mundo en que vivimos se trae a un primer plano en lugar de ser pre-dado, la noción de representación ya no puede cumplir un rol central. La profundidad de las suposiciones que estamos tocando aquí no debe subestimarse, ya que nuestra tradición racionalista totalizadora ha favorecido (con variantes, por supuesto) la comprensión del conocimiento como un espejo de la naturaleza. Es solamente en el trabajo de ciertos pensadores continentales (principalmente, M. Heidegger, M. Merleau-Ponty y M. Foucault), que ha comenzado una crítica explícita a las representaciones y la dimensión enactiva del conocimiento. Estos temas hermenéuticos se introdujeron por primera vez como la disciplina de la interpretación de textos antiguos, pero en la actualidad sus alcances se han extendido para denotar el fenómeno completo de la *interpretación* entendida como la actividad de la puesta en activación o "traer a un primer plano" que ya hemos

mencionado[34]. Dado que lo que nos ocupa aquí es la predominancia del uso, en lugar de las representaciones, nos parece adecuado llamar enactivo a este enfoque alternativo a la ciencia cognitiva[35].

En años recientes, sin embargo, unos pocos investigadores dentro de la ciencia cognitiva han presentado propuestas concretas, llevando esta crítica desde el nivel filosófico al laboratorio y al trabajo específico en la Inteligencia Artificial. Esto constituye un alejamiento más radical de la ciencia cognitiva que el precedente, y que va más allá de los temas discutidos durante el período formativo. Al mismo tiempo incorpora de manera natural las ideas y métodos desarrollados en el ámbito conexionista, como veremos ahora.

El problema con la resolución de problemas

En la ciencia cognitiva siempre se ha supuesto que el mundo puede dividirse en una región de elementos discretos y tareas a las cuales se dirige el sistema cognitivo mismo, actuando dentro de un 'dominio' de problemas: visión, lenguaje, movimiento. Aunque es relativamente fácil definir todos los posibles estados en el 'dominio' del juego de ajedrez, ha probado ser menos productivo trasladar este enfoque a, digamos, el 'dominio' de los robots móviles. Por cierto que uno también puede aquí individualizar ítemes discretos (como marcos de acero, ruedas y ventanas en el ensamble de un auto). Pero también es claro que mientras el mundo del ajedrez termina ordenadamente en algún punto, el mundo del movimiento entre los objetos no. Se requiere un permanente uso del sentido común de nuestra parte para configurar nuestro mundo de objetos.

---

34 De la mayor influencia en este respecto es el trabajo de H.G. Gadamer, *Truth and Method*, Seabury Press, 1975. Para una clara introducción a la hermenéutica, ver Palmer, *Hermeneutics*, Northwestern University Press, 1979. La formulación de esta sección debe mucho a la influencia de F. Flores: ver T. Winnogard y F. Flores *Understanding Computers and Cognition: A New Foundation for Design*, Ablex, Nueva Jersey, 1986.

35 El nombre dista mucho de estar establecido. Lo sugiero aquí por razones pedagógicas hasta que se proponga uno nuevo.

De hecho, lo que es interesante acerca del sentido común es que no puede ser incorporado al conocimiento en absoluto, ya que es más bien una facilidad o un saber basado en la experiencia vivida y un vasto número de casos, lo cual implica la encarnación de una historia. Un estudio cuidadoso de la adquisición de habilidades, por ejemplo, parece confirmar este punto[36]. Un mundo vivido, natural, no tiene fronteras nítidas y, por lo tanto, esperamos que una representación simbólica con normas no sea capaz de captar la cognición fundada en el sentido común. De hecho, es justo decir que para la década de los 70, después de dos décadas de avance degradantemente lento, muchos trabajadores de las ciencias cognitivas se dieron cuenta de que incluso la acción cognitiva más simple requiere una cantidad aparentemente infinita de conocimiento, que damos por sentado y obvio, pero que debe ser dado con cuchara a la computadora. La esperanza que tenían a principios de los sesenta los cognitivistas de encontrar una máquina que resolviera los problemas de carácter general debió reducirse a dominios de conocimiento local, con problemas por resolver claramente planteados, en los que el programador pudiera entregar a la máquina tanto como fuera posible de su propio conocimiento general. De manera similar, la estrategia conexionista comúnmente puesta en práctica depende de una restricción del ámbito de posibles atractores mediante la aplicación de suposiciones acerca de propiedades conocidas del mundo que se incorporan como restricciones adicionales a la regularización[37] o, más recientemente, en métodos de propagación a posteriori, como un modelo perfecto para ser imitado. En ambas instancias, la inmanejable ambigüedad del sentido común contextual se deja en la periferia de la investigación, con la esperanza de que se aclarará a su debido tiempo[38].

---

[36] H. Dreyfus y S. Dreyfus, *Mind over Machine*, Free Press /MacMillan, Nueva York, 1986.

[37] Para leer sobre esta explícita manera de construir redes de inspiración biológica, ver T. Poggio, V. Torra y C. Koch, *Nature* 317, 314-319 (1986).

[38] Para un interesante ejemplo de una discusión en Inteligencia Artificial acerca de estos temas, ver el análisis múltiple del libro de Winnograd y Flores en *Artif. Intell.* (1987).

Estas preocupaciones admitidas tienen una contraparte filosófica bien desarrollada. Los Fenomenólogos de tradición continental han provocado detalladas discusiones acerca de por qué el conocimiento es materia del ser en un mundo que es inseparable de nuestros cuerpos, nuestro lenguaje y nuestra historia social[39]. Es una interpretación de carácter continuo que no puede fijarse adecuadamente en un conjunto de normas y suposiciones dado que es materia de acción y de historia, una cognición recogida por imitación y que pasa a ser parte de una cognición preexistente. Más aún, no podemos pretender pararnos afuera del mundo en el que nos encontramos para reflexionar acerca de la manera en que sus contenidos coinciden con sus representaciones del mismo: siempre habremos estado ya inmersos en él. Considerar las normas como actividad mental equivale a desarticular la coyuntura misma de donde surge la cualidad viva de la cognición. Esto sólo podría hacerse en un contexto muy limitado en donde casi todo fuera constante, una condición *ceteris paribus* generalizada. Contexto y sentido común no son elementos residuales que puedan eliminarse de manera progresiva mediante el descubrimiento de normas más sofisticadas. Son de hecho la esencia misma de la cognición *creativa*.

Si esta crítica es correcta, aunque sea sólo hasta cierto punto, el avance en la comprensión de la cognición tal como funciona normalmente (y no exclusivamente en ambientes altamente restringidos) no se producirá, salvo que partamos de una base distinta a la de un dominio que existe ahí afuera para ser representado.

Las representaciones abandonan el escenario

El verdadero desafío que implica para la ciencia cognitiva esta orientación es el cuestionamiento de la más arraigada suposición de nuestra tradición científica: que el mundo como lo experimentamos es independiente del conocedor. En lugar de eso, si somos

---

[39] Los principales puntos de referencia que tenemos en mente en este punto son(en sus versiones en Inglés): M. Heidegger, *Being and Time*, Harper and Row, 1977; M. Merleau-Ponty, *The Phenomenology of Perception*, Routlege y Kegan Paul, 1962; Michel Foucault, *Discipline and Punish: the birth of the prison*, Vintage/Random House, 1979.

forzados a concluir que la cognición no puede entenderse apropiadamente sin el sentido común, el cual no es sino nuestra *historia corporal y social*, la inevitable conclusión es que el conocedor y lo conocido, sujeto y objeto, están en una relación de mutua especificación: surgen conjuntamente.

Consideremos el caso de la visión: ¿qué vino primero, el mundo o la imagen? La respuesta de la investigación de la visión (tanto cognitivista como conexionista) surge de manera clara a partir de los nombres de las tareas investigadas: 'recobrar la forma a partir de las sombras' o 'la profundidad a partir del movimiento' o 'el color a partir de fuentes de iluminación variables'. Podemos llamar a esto el extremo del huevo o la gallina:

- Postura de la *Gallina*: El mundo afuera de nosotros tiene leyes fijas, precede a la imagen que proyecta sobre el sistema cognitivo, cuya tarea es capturar esta imagen apropiadamente (ya sea como símbolos o en estados emergentes).

Ahora, por favor tomen nota de lo razonable que esto suena y lo difícil que parece imaginar que pudiera ser de otra manera. Tendemos a pensar que la única alternativa es la postura del *Huevo*:

- Postura del *Huevo*: el sistema cognitivo crea su propio mundo, y toda su aparente solidez es la reflexión primaria de las leyes internas del organismo.

La orientación enactiva propone que tomemos una vía intermedia,[40] yendo más allá de estos dos extremos, dándonos cuenta (como lo saben los granjeros) que el huevo y la gallina *se definen el uno al otro*, son correlativos. Es el permanente proceso de vivir lo que ha dado forma a nuestro mundo en el ir y venir entre lo que describimos como restricciones externas, desde nuestro punto de vista perceptual, y la actividad generada internamente. Los orígenes de este proceso se han perdido para siempre y nuestro mundo, para todos los efectos prácticos, es estable (... excepto cuando se des-

---

[40] Esto se discute en mi contribución al anterior Symposium Stanford, 'Living ways of sense making: A middle way approach to neuroscience' en P. Livingston (Ed.) *Order and Disorder*, Anma Libris, Stanford, 1984.

compone). Pero esta aparente estabilidad no necesariamente obscurece la búsqueda del mecanismo que los hizo surgir. Es este énfasis en una co-determinación (más allá de las posturas del huevo y la gallina) lo que marca la diferencia entre el punto de vista enactivo y cualquier forma de constructivismo[41] o neo-kantismo biológico[42]. Es importante tener esto presente, ya que la filosofía más o menos realista que satura a la ciencia cognitivista tenderá a asumir que cualquiera que cuestione las representaciones deberá *ipso facto* estar en una posición diametralmente opuesta, en la que el fantasma del solipsismo también existe.

Color y olor como ejemplos

Las consideraciones anteriores por lo general se hacen en el ámbito del lenguaje y las comunicaciones humanas, y parecería ser que para el mundo perceptual más inmediato, no serían relevantes. Sin embargo, el punto es que la Enacción se aplica en *todos* los niveles y por lo tanto es importante examinar la percepción bajo este punto de vista.

Considere el mundo de colores que percibimos diariamente. Normalmente se asume que el color es un atributo de la longitud de onda de la luz reflejada por los objetos, el cual recogemos y procesamos como información relevante. De hecho, como se ha documentado extensamente en la actualidad, el color percibido de un objeto es en gran medida independiente de la longitud de onda entrante[43]. En lugar de eso, existe un complejo (y sólo parcialmente comprendido) proceso de comparación cooperativa entre múltiples conjuntos neuronales en el cerebro, el cual especifica el color de un objeto de acuerdo con el estado global que alcanza: se especifica un espacio perceptual cromático[44].

---

41 Ver, por ejemplo. P. Watzlawick (Editor), *The Invented Reality: Essays on Constructivism*, Norton, Nueva York, 1985.

42 Esto se ve más claramente en la escuela de Viena de Konrad Lorenz, como se expresa, por ejemplo, en *Behind the Mirror*, Harper and Row, 1979.

43 Ver, por ejemplo, E. Land, *Proc. Natl. Acad. Sci.* (USA) **80**, 5163-5169 (1983).

44 P. Gouras y E. Zenner, *Progr. Sensory Physiol.* **1**, 139-179 (1981)

Ahora bien, es claro que estos mecanismos son coherentes con lo que describimos como restricciones a la iluminación (reflectancia, discontinuidad de los objetos, y así sucesivamente) pero no son una consecuencia *lógica* de las mismas. Las operaciones neuronales cooperativas que subyacen a nuestra percepción del color han sido el resultado de la larga evolución biológica del grupo de los primates. Sus efectos han penetrado a tal grado nuestra vida que resulta tentador asumir que los colores, tal cual los vemos, son como el mundo *es*. Pero esta conclusión se modera si recordamos que otras especies, en su evolución, han gestado distintos mundos cromáticos mediante operaciones neuronales cooperativas propias de sus órganos sensoriales. Por ejemplo, el espacio cromático de la paloma es aparentemente tretracromático (requiere cuatro colores primarios) a diferencia de nosotros los tricrómatas (para quienes son suficientes sólo tres colores primarios)[45]. Esto no es meramente una expansión de la diversidad en un mismo espectro, sino una dimensión enteramente nueva que da lugar a un mundo cromático tan inconmensurable al nuestro como el nuestro lo es para una persona daltónica. Aquí, el color aparece no como un correlato de las propiedades del mundo, sino como regularidades que son co-definidas con una modalidad de ser particular.

Lo que sí puede decirse es que nuestro mundo cromático es *viable*: es efectivo dado que hemos continuado nuestro linaje biológico. Las historias vastamente diferentes de acoplamiento estructural de las aves, insectos y primates han dado lugar a un mundo relevante para cada especie que es inseparable de su vivir. Todo lo que se necesita es que cada camino tomado sea viable, es decir, que constituya una serie ininterrumpida de cambios estructurales. Los mecanismos neuronales subyacentes al color no son la solución a un 'problema (recoger las propiedades cromáticas correctas de los objetos), sino que surgen conjuntamente de la percepción del color *y* de lo que uno puede describir como atributos cromáticos del mundo en que habita.

---

[45] F. Varela et. al. *Arch. Biol. Med. Exp.* 16, 291-303 (1983); E. Thompson, A. Palacios, F. Varela *Beh. Brain. Sci.*(en imprenta), 1992.

Otra dimensión perceptual en que estas ideas pueden verse en juego es el olfato, no a causa de los alcances comparativos proporcionados por la filogenia, sino a las novedosas técnicas electro-fisiológicas. Luego de muchos años de trabajo, Freeman[46] ha logrado insertar un gran número de electrodos en el abultamiento olfatorio/órgano del olfato de un conejo, de tal manera que una pequeña porción de la actividad global puede medirse mientras el animal se mueve libremente. Se encontró que no existe un patrón claro de actividad global en el bulbo, salvo que el animal se vea expuesto a un olor específico varias veces. Más aún, dichos patrones emergentes parecen crearse a partir de un contexto de actividad incoherente en un atractor coherente. Al igual que en el caso del color, el olor se revela a sí mismo no como una cartografía de las características externas, sino como el dimensionamiento creativo del significado sobre la base de la historia[47].

Entonces, a la luz de lo anterior, el funcionamiento del cerebro está centralmente preocupado con la constante escenificación de mundos a través de la historia de los linajes viables; un órgano que hace mundos en lugar de reflejarlos.

Un resumen de la doctrina

Entonces, el concepto básico es que las capacidades cognitivas están inseparablemente unidas a una historia que es vivida, algo muy parecido a un camino que no existe, pero que se hace al andar. En consecuencia, el punto de vista de la cognición no es la resolución de problemas a través de representaciones, sino el traer a un primer plano, creativamente, a un mundo en que la única condición exigida es que constituya una *acción efectiva*: que permita la continuidad de la integridad del sistema involucrado[48].

---

46 W. Freeman, *Mass Action in the Nervous System*, Academic Press, 1975

47 W. Freeman y C. Skarda, *Brain Res. Reviews*, 10, 145-175 (1985). Significativamente, una sección de este artículo se titula: 'A Retraction on "representation"' (página 169).

48 Esta reinterpretación biológicamente inspirada del entendimiento fue presentada en H. Maturana y F. Varela, *Autopoiesis and Cognition: The Realization of the Living*, D. Reidel. Boston, 1980, y F. Varela, *Principles of Biological Autonomy*, op cit. Para una exposición

Pregunta 1: ¿Que es la cognición?

Respuesta: Acción efectiva: Una historia de acoplamiento estructural que escenifica (genera) un mundo.

Pregunta 2: ¿Cómo funciona?

Respuesta: A través de una red de elementos interconectados capaces de cambios estructurales que experimentan una historia ininterrumpida.

Pregunta 3: ¿Cómo puedo saber cuándo un sistema cognitivo está funcionando adecuadamente?

Respuesta: Cuando se vuelve parte de un mundo de significados existente y contínuo (en la ontogenia) o forma uno nuevo (en la filogenia).

Debe tomarse nota de que a partir de estas respuestas surgen dos ideas que normalmente no se consideran en la ciencia cognitiva. Una de ellas es que, debido a que las representaciones no juegan un papel central, la inteligencia ha pasado de ser la capacidad para resolver un problema a ser la capacidad de *ingresar* a un mundo compartido. La segunda idea es que lo que toma el lugar del diseño orientado al cumplimiento de una tarea es un proceso *evolutivo*. Expresado de manera burda, así como el conexionismo surge a partir del cognitivismo inspirado por un contacto más cercano con el cerebro, la orientación enactiva da un paso más adelante en la misma dirección para abarcar la temporalidad de vivir ya sea en la ontogenia o la filogenia.

Trabajar sin representaciones

La búsqueda de alternativas a la representación para el estudio de fenómenos cognitivos (lo cual, admitámoslo, es un concepto englobador algo vago, tal como lo es el conexionismo) atrae a

---

introductoria de este punto de vista y desarrollos más recientes, ver. H. Maturana y F. Varela, *The Tree of Knowledge: the Biological Roots of Human Understanding*. New Science Library, Boston, 1987. Las conexiones con el lenguaje y la Inteligencia Artificial se discuten en Winnograd and Flores, *op. cit.*

un grupo relativamente pequeño de personas en diversos campos. Más aún, como argumentaré en la próxima sección, muchas de las herramientas utilizadas en la perspectiva conexionista tradicional pueden reformularse en este contexto, de tal manera que las líneas divisorias son mucho menos nítidas aquí de lo que fueron entre las orientaciones simbólica y conexionista.

Es claro que una estrategia enactiva para la Inteligencia Artificial sólo es factible si uno está dispuesto a relajar las restricciones de una operatoria específica en la resolución de problemas. Este es el espíritu, por ejemplo, de los llamados sistemas clasificadores,[49] concebidos para confrontar un medio ambiente no definido al cual se debe dar forma y significado. En términos más generales, las simulaciones de historias prolongadas de acoplamientos con diversas estrategias evolutivas permiten descubrir tendencias al interior de las cuales surgen desempeños cognocitivos[50]. Pero estas nuevas perspectivas para la investigación se encuentran solamente en sus inicios más tempranos.

7. Vinculación de la emergencia y la enacción

El vínculo entre la emergencia y la enacción depende de la transformación de nuestra lectura de lo que puede hacer un sistema distribuido. Si uno pone más enfasis en la forma en que un proceso histórico conduce a regularidades emergentes sin una restricción final fija, uno recupera la condición biológica más abierta. Por el contrario, si uno pone énfasis en cómo una red adquirirá una capacidad muy específica en un dominio muy definido (por ejemplo en NetTalk),

---

49 Ver J.H. Holland, 'Escaping brittleness', en: *Machine Intelligence*, Vol 2 (1986).

50 Una interesante colección reciente de diversos escritos en esta dirección puede encontrarse en: *Evolution, Games and Learning: Models For Adaptation in Machines and Nature, Physical 22D* (1986). Con seguridad, muchos de los contribuyentes no estarían de acuerdo con nuestra lectura de sus trabajos. Para obtener un ejemplo explícito, ver: F. Varela 'Structural coupling and the origin of meaning in a simple cellular automata' in E. Secarz (Ed.) *The Semiotics of Cellular Communications*, Springer-Verlag, Nueva York, 1987.

entonces se vuelve a las representaciones y tenemos el punto de vista más común acerca de los modelos conexionistas. Sin embargo, la primera interpretación también implica una perspectiva totalmente nueva sobre lo que es la cognición, como se resume en la sección anterior.

De este modo, el camino tomado depende, en gran medida, del grado de interés que se tenga en permanecer cerca de la realidad biológica y más alejado de consideraciones pragmáticas y de ingeniería. Por supuesto que es posible definir un dominio fijo dentro del cual un sistema conexionista puede funcionar, pero con ello se oscurecen los temas más profundos acerca del origen, centrales para el punto de vista enactivo.

Consideremos, por ejemplo, la teoría de la armonía de Smolesnky. Su punto de vista de la computación sub-simbólica como un modelo para la intuición parece estar eminentemente en línea con una perspectiva enactiva, razón por la cual puede servir como el mejor caso para considerar aquí con fines de contraste. Sin embargo, incluso la teoría de la armonía se evalúa en referencia a un nivel inviolado de realidad medioambiental: las características exógenas coinciden con características dadas del mundo y la actividad endógena que adquiere a través de la experiencia un estado abstracto de significación que "codificará óptimamente la regularidad medioambiental". La esperanza es encontrar una actividad endógena que corresponda a una "caracterización del entorno basada en la optimalidad"[51]. El punto de vista enactivo requeriría llevar este tipo de sistema cognitivo a una situación en la que lo endógeno y lo exógeno fueran mutuamente definitorios a través de una prolongada historia que requiriera únicamente un acoplamiento viable, renunciando a cualquier forma de calce óptimo[52].

Es cierto que, desde el punto de vista de una Inteligencia Artificial de orientación pragmática, que tenga como objetivo la generación de un sistema que opere en algún dominio en el corto plazo, esta orien-

---

[51] P. Smolesnky, *op.cit.* p. 260.

[52] Vale la pena notar que es posible aplicar argumentos similares al pensamiento evolucionista en la actualidad. Para obtener paralelos entre el representacionismo cognitivista y el adaptacionismo evolucionista, ver F. Varela, en: P. Livingston (Ed.), *op.cit* y la Introducción en este volumen.

tación parece no tener sentido. Mi argumento es que las propiedades cognitivas emergieron en los sistemas vivos sin esas consideraciones de optimalidad. Son el resultado de historias de compensaciones viables que crean regularidades, pero dista mucho de ser evidente que pueda decirse que corresponden a algún referente único.

Por lo tanto, existe una tensión entre los dos mundos paralelos de investigación, en que la opción de estar a favor o en contra de la crítica enactiva se toma de acuerdo con todas las complejidades de un giro conceptual por un lado, y por otro el mundo tecnológico, en donde la camisa de fuerza de la aplicabilidad inmediata fija los límites de su extensión. Me parece que esta tensión probablemente se resolverá a través de una creciente distancia entre los componentes tecnológicos y científicos de la ciencia cognitiva.

8. Conclusión: conocimiento materializado

Comenzamos con la línea dura de la ciencia cognitiva y nos desplazamos hacia lo que podría considerarse su periferia, esto es, la consideración del contexto circundante y los efectos de la historia biológica y cultural sobre la cognición y la acción. Por cierto que aquellos que siguen adhiriendo a las representaciones como una idea clave, ven estos aspectos como sólo temporalmente externos al ámbito más preciso de la orientación resolutoria de problemas, la cual parece más accesible. Otros van al extremo de adoptar la postura de que tales aspectos 'borrosos' y 'filosóficos' ni siquiera deberían formar parte de una ciencia cognitiva propiamente tal.

Algunos de los contrastes que generan estas tensiones pueden estar representados en la siguiente tabla:

Desde	Hacia
específico para una tarea	creativo
eesolución de problemas	definición de problema
abstracto, simbólico	historia, vinculado al cuerpo
universal	sensible al contexto
centralizado	distribuido
secuencial, jerárquico	paralelo

mundo pre-dado representación implementación por diseño	mundo gestado acción efectiva implementación por estrategias evolutivas
abstracto	*materializado*

A modo de esquema visual de esta presentación, he resumido los tres principales enfoques discutidos en este texto en el *mapa polar* de la Figura 2.

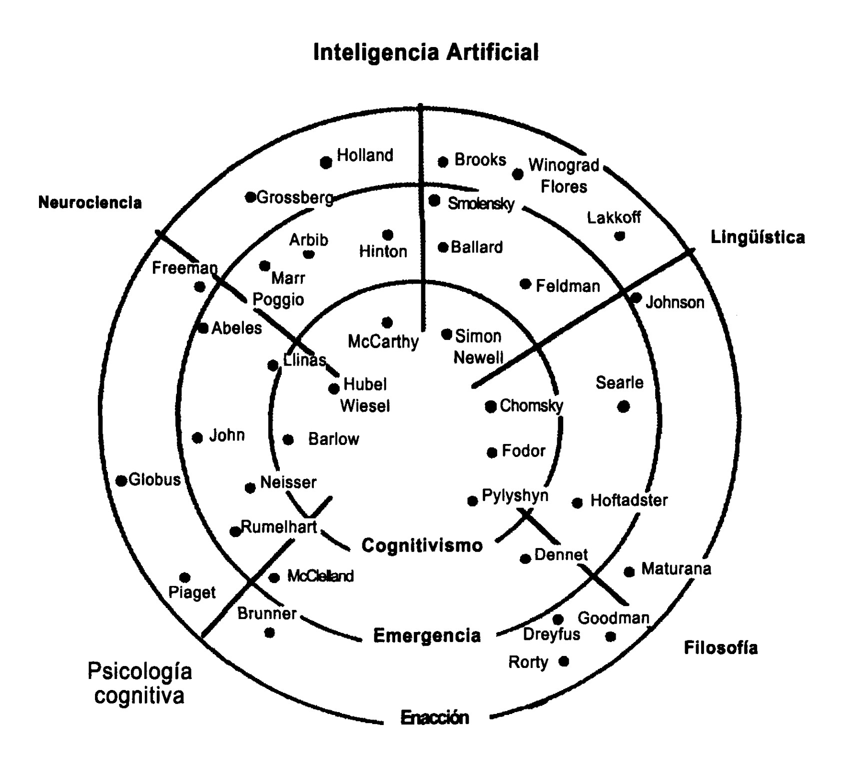


Figura 2: Un mapa polar de STC, con el paradigma cognitivista al centro y los temas alternativos formando una periferia, ambos tocando el cuerpo de doble filo de las ideas conexionistas. A lo largo de los radios disciplinarios aparecen los nombres de algunas personas que trabajan en cada área representativa.

Mi punto de vista es que estas tres olas sucesivas para la comprensión de la cognición básica y su origen se relacionan entre sí por una *imbricación* sucesiva, como cajas chinas. En la dirección centrípeta, una va desde la emergencia a lo simbólico, estableciendo categorías en la base a partir de la cual emergen los símbolos, y trabajando con los símbolos en su valor nominal. O bien uno puede avanzar desde la enacción hasta una visión conexionista estándar asumiendo ciertas regularidades dadas del dominio en el cual opera el sistema (es decir, una función de idoneidad en un dominio). La dirección centrífuga es una categorización progresiva de lo que parece estable y regular, para considerar de donde pueden haber surgido dichas regularidades, hasta e incluyendo las dimensiones perceptuales de nuestro mundo humano.

Es claro que cada uno de estos enfoques, como niveles de descripción, es útil en su propio contexto. Sin embargo, si nuestra tarea es comprender el origen de la percepción y la cognición tal como las encontramos en nuestra historia efectivamente vivida, pienso que el nivel correcto de explicación es el borde exterior más inclusivo del mapa. Además, para una Inteligencia Artificial en la que las máquinas sean inteligentes en el sentido de desarrollarse hacia un sentido común compartido con los seres humanos, como pueden hacerlo los animales, no veo otro camino que hacerlas pasar por un proceso de transformaciones evolutivas como se sugiere en la perspectiva enactiva[53]. Qué tan fértil, difícil o imposible probará ser esto, es por ahora una mera adivinanza.

Mis preferencias han sido expresadas explícitamente en el texto. En especial, en este ensayo he alegado que si el pivote articulador de la cognición es su capacidad para generar significados, entonces la información no está preestablecida como un orden dado, sino que está constituida por las regularidades que emergen

---

[53] Esta es la tendencia dentro del nuevo campo de la 'Vida Artificial'; ver por ejemplo Ch. Langton (Ed.) *Artificial Life*, Addison-Wesley, Nueva Jersey, 1990.

de las actividades cognitivas mismas. Es este re-enmarcado lo que tiene múltiples consecuencias éticas, que a estas alturas deberían ser evidentes. Como se indica en el último ítem de la tabla antes indicada, en la medida en que nos desplazamos desde una visión abstracta a una visión plenamente materializada del conocimiento, los hechos y los valores se vuelven *inseparables*. Saber *es* evaluar a través de nuestro vivir, en una circularidad creativa.

EL REENCANTAMIENTO DE LO CONCRETO

El tema central aquí es el paso entre el origen del significado y su relación con la vida encarnada momento a momento. Es lo que presento aquí bajo la noción de micromundos y su fragilidad intrínseca. El texto sirve de preludio y transición a la importancia de la fenomenología, que es lo central de la próxima Sección.

1. El desencantamiento de lo abstracto

Virajes en las ciencias cognitivas. "Racionalista", "cartesiana" u "objetivista": estos son algunos de los términos que se utilizan hoy en día para caracterizar la tradición dominante en la que hemos crecido[1]. Sin embargo, cuando se trata de reformular el conocimiento y la cognición, considero que el término que mejor se adecua a nuestra tradición es "abstracta": no hay palabra que caracterice mejor a las unidades de conocimiento consideradas más "naturales". Esta tendencia para abrirnos paso hasta la atmósfera enrarecida de lo general y lo formal, de lo lógico y lo bien definido, de lo representado y lo planificado, es lo que hace que nuestro mundo occidental nos sea tan familiar.

(Varela, F. (1995), The re-enchantement of the concrete, in: S.Kwinter and J.Cleary (Eds.), *Incorporations*, Zone Books, New York, pp.320-340, 1992.)

[1] Fragmentos de este texto aparecen, con algunas modificaciones, en James Ogilvy, ed., *Revisioning Philosophy* (New York: SUNY Press, 1991) y en Jacques Montangero, ed., *Psychologie génétique et sciences cognitives*, Cahiers de la Fondation Jean Piaget nº 11 (Genève, 1991).

La tesis que quisiera plantear en este trabajo es la siguiente: existen claras indicaciones de que el conjunto de ciencias que tratan del conocimiento y de la cognición —las ciencias cognitivas— lentamente han ido cobrando conciencia de que las cosas han sido planteadas al revés, y han comenzado un radical viraje paradigmático o epistémico. El núcleo de esta visión emergente es la convicción de que las verdaderas unidades de conocimiento son de naturaleza eminentemente concreta, encarnada, incorporada, vivida. Este conocimiento concreto, su historicidad y su contexto, no es un "ruido" que oscurece la pureza de un esquema que ha de ser captado en su verdadera esencia, una configuración abstracta, como tampoco es un paso hacia otra cosa: es cómo llegamos y dónde permanecemos.

Tal vez lo que mejor nos permita ilustrar esta tendencia es la transformación gradual de las ideas en el campo eminentemente pragmático de la inteligencia artificial. Durante sus dos primeras décadas (1950-1970), la investigación se basó en el paradigma computacionalista, según el cual el conocimiento opera de acuerdo a reglas de tipo lógico en la manipulación simbólica, idea que encuentra toda su expresión en los modernos sistemas digitales. Inicialmente, los esfuerzos se orientaron a la resolución de problemas de índole general, tales como la traducción del lenguaje natural o la creación de un "aparato para la resolución general de problemas". Estos intentos, mediante los cuales se buscaba igualar la inteligencia de un experto altamente capacitado, fueron considerados la mejor forma de abordar los problemas más importantes de la cognición. A medida que los esfuerzos por alcanzar estos objetivos iban fracasando, quedó en evidencia que la única forma de lograr algún adelanto era introducir pretensiones más modestas, ya que las tareas más sencillas, incluso aquellas realizadas por insectos pequeños, son sencillamente imposibles de llevar a cabo mediante una estrategia computacional. Es así como a este período inicial siguió la convicción de que había que invertir al experto y al niño en la escala de rendimiento: quedó en evidencia que la inteligencia más fundamental es la del bebé que adquiere el lenguaje a partir de elocuciones sueltas, y que es capaz de crear objetos significativos a partir de un mundo previamente inespecífico.

Esta perspectiva ha permitido rescatar el papel de lo concreto, focalizando a escala apropiada la actividad cognitiva que se desarrolla en un espacio muy peculiar, que podemos llamar los "goznes" del presente inmediato. Porque es en el present*e inmediato* que vive realmente lo concreto. Pero antes de proceder, es necesario revisar algunos de los supuestos más arraigados que hemos heredado de la ortodoxia computacional.

Acerca de los agentes cognitivos des-unidos. Hay considerable evidencia de que los cerebros no son máquinas lógicas, sino redes sumamente cooperativas, no homogéneas y distribuidas. Podríamos decir que todo el sistema se asemeja a un *mosaico* de subredes, ensamblado por una compleja historia de transformaciones, más que un sistema optimizado producto de un diseño limpio y unificado. Este tipo de arquitectura también sugiere que, en vez de buscar grandiosos modelos unificados de todos las conductas en red, habría que estudiar aquellas redes cuyas habilidades cognitivas están restringidas a actividades específicas y concretas mientras interactúan entre sí.

Los científicos cognitivistas han comenzado a tomar en serio esta perspectiva de la arquitectura cognitiva. Por ejemplo, según Marvin Minsky, la mente estaría constituida por una sociedad de "agentes" cuyas habilidades están bastante circunscritas: cada agente tomado en forma individual trabaja solamente con problemas a pequeña escala o "de juguete"[2] ya que los problemas se vuelven inmanejables para una sola red a medida que aumentan proporcionalmente (esto es algo que los científicos cognitivistas sólo entendieron hace algún tiempo). La tarea consiste en organizar a los "agentes" que operan en estos dominios específicos, en sistemas o "agencias" de mayor tamaño y eficacia para luego transformar estas agencias en sistemas de un nivel más alto. De este modo, la mente sería equivalente a un tipo de "sociedad".

Es importante recordar que si bien se inspira en una nueva forma de concebir el cerebro, éste es un modelo de la *mente.* En

---

[2] Marvin Minsky, *The Society of Mind* (Nueva York: Simon and Schuster, 1986).

otras palabras, no es un modelo de sociedades neurales o de redes: es un modelo de la arquitectura cognitiva que hace caso omiso de lo neurológico, que se sustrae de la experiencia viva y vivida. Por lo tanto, los agentes y las agencias no son entidades ni procesos materiales: son procesos abstractos o funciones. Es importante destacar este punto, sobre todo porque hay momentos en que Minsky escribe como si estuviera hablando de la cognición a nivel del cerebro. Como señalaré más adelante, lo que falta es el vínculo entre estos agentes y el acoplamiento encarnado, a través de la sensación y la acción, que es inherente a la cognición viva. Nos detendremos ahora algunos instantes para examinar algunas de las implicaciones de esta perspectiva de subredes cognitivas fragmentadas y locales.

El modelo de la mente concebido como una sociedad compuesta por numerosos agentes intenta abarcar diferentes enfoques sobre la cognición, desde la noción de redes distribuidas autoorganizadoras hasta la clásica perspectiva cognitivista de un procesamiento simbólico. Esta perspectiva representa un desafío para el modelo de la mente centralizado o unificado, tanto bajo la forma de redes distribuidas como de procesos simbólicos. Esto es evidente, por ejemplo, cuando Minsky argumenta que hay virtudes no sólo en la distribución, sino también en la insulación, esto es, en los mecanismos que mantienen separados a los diferentes procesos. Esta idea también ha sido extensamente explorada, aunque en un contexto diferente, por Jerry Fodor[3]. Los agentes de una agencia pueden estar conectados bajo la forma de una red distribuida, pero si las agencias estuvieran conectadas de este mismo modo, constituirían una gran red cuyas funciones estarían uniformemente distribuidas. Sin embargo, esta uniformidad restringiría la capacidad de las agencias individuales para combinar sus operaciones en forma productiva. Cuanto más distribuidas las operaciones, tanto más difícil les resulta a las agencias actuar al mismo tiempo sin interferencia mutua. Sin embargo, si hay mecanismos que mantienen aisladas a las agencias, estos problemas

---

[3] Jerry Fodor, *The Modularity of Mind* (Cambridge, Mass.: Bradford Books and MIT Press, 1983).

se evitan, y las agencias seguirían interactuando, pero a través de conexiones más limitadas.

Claro está que los detalles de este tipo de enfoque programático son debatibles, pero el esquema general sugiere (lo que no sólo es propio de la formulación de Minsky en términos de agentes y agencias) que la mente no es una entidad homogénea y unificada, y ni siquiera un conjunto de entidades, sino u*n conjunto heterogéneo de procesos desunificados*. Este conjunto des-unificado obviamente puede ser examinado desde diferentes niveles. Si cambiamos de perspectiva, lo que cuenta como una agencia (esto es, un grupo de agentes), simplemente podría ser considerado como un agente en una agencia más grande. Y a la inversa, si lo enfocamos en más detalle, lo que cuenta como un agente podría ser considerado como una agencia compuesta por muchos agentes. Del mismo modo, lo que cuenta como una sociedad, también dependerá de la perspectiva que escojamos.

Habiendo preparado la escena para este problema central de las ciencias cognitivas contemporáneas, quisiera abordar algunas de sus implicaciones para el tema que nos ocupa: cómo lo propio de lo concreto es su estar centrado en el presente.

2. El estar ahí: Durante los quiebres

Disposición a la acción en el presente. Mi preocupación en este momento tiene que ver con una de las muchas consecuencias que plantea esta perspectiva de la des-unidad del sujeto, entendido como agente cognitivo. La pregunta puede ser formulada del siguiente modo: dado que en cada acto cognitivo hay una multitud de subprocesos que compiten, ¿cómo debemos comprender el momento de negociación y emergencia en que surge uno de estos subprocesos y constituye una conducta definida? Dicho de manera más elocuente: ¿cómo debemos entender el preciso momento de estar ahí en que aparece algo concreto y específico?

Imagine que va caminando por la calle, yendo a reunirse con alguien. La jornada está a punto de concluir y su mente no está

ocupada en nada muy especial. Se encuentra relajado, en lo que podríamos llamar la disposición de una persona que da un paseo. Introduce la mano en el bolsillo y de pronto descubre que su billetera no se encuentra en el lugar habitual. Quiebre: se detiene, su estado mental es confuso, su tonalidad emocional cambia. Antes de que se percate de ello, surge un mundo diferente. Intuye claramente que dejó su billetera en la tienda donde acaba de comprar un paquete de cigarrillos. Ahora su estado anímico se transforma en preocupación por la pérdida de documentos y de dinero, la disposición a la acción es regresar rápidamente a la tienda. Presta poca atención a los árboles que lo rodean y a las personas que pasan. Toda su atención está centrada en perder el menor tiempo posible.

Situaciones como ésta constituyen la trama de nuestras vidas. Siempre operamos en la inmediatez de una situación dada. Nuestro mundo vivido está tan a la mano que no tenemos un control deliberado acerca de lo que es y de cómo lo habitamos. Cuando nos sentamos a la mesa para comer con un amigo o un pariente, el complejo *know-how* que implica el manejo de los utensilios, las posturas corporales y las pausas en la conversación están todas presentes sin que exista deliberación. Nuestro yo-almorzando es transparente[4]. Terminamos de almorzar, regresamos a la oficina y habitamos una nueva disposición con un modo diferente de hablar, un tono postural diferente y diferentes evaluaciones. Tenemos una disposición a la acción propia de cada situación específica que vivimos. Nuevas formas de conducta y las transiciones y puntuaciones que las separan corresponden a miniquiebres que experimentamos continuamente. Hay ocasiones en que los quiebres no son tan "mini", sino más bien macroscópicos, como cuando vivimos un peligro repentino. Me referiré a esta disposición a la acción como "microidentidad", y a la situación correspondiente como "micromundo". La forma *como* nos

[4] La noción de transparencia ha sido extensamente desarrollada en un manuscrito no publicado de Fernando Flores y Michael Groves (Logonet, Inc., Berkeley, Calif., 1990). Le agradezco al Sr. Flores por haberme permitido leer este trabajo del que mis propias ideas se han beneficiado.

presentamos no puede disociarse de la forma en que las cosas y los demás se presentan *ante* nosotros. Aquí podría incursionar un poco en la fenomenología e identificar algunos de los micromundos típicos en que nos movemos en un día normal. No se trata, sin embargo, de catalogarlos, sino de reconocer su *recurrencia*: ser capaz de una acción efectiva es, en un sentido importante, la forma en que encarnamos un conjunto de transiciones recurrentes entre micromundos. No estoy diciendo que no existan situaciones en que la recurrencia no es aplicable. Por ejemplo, cuando llegamos por primera vez a un país extranjero existe una enorme ausencia de conductas a la mano y de micromundos recurrentes, Muchos actos sencillos, como la conversación social o el comer, tienen que realizarse en forma deliberada o aprenderse. En otras palabras, los micromundos y las microidentidades se constituyen históricamente. Pero la forma de vida más común consiste en micromundos ya constituidos y que componen nuestra identidades. Claramente, todavía queda mucho por investigar y decir sobre la fenomenología de la experiencia común[5]. Mi intención aquí es más modesta: simplemente destacar una serie de fenómenos muy cercanos a nuestra experiencia común.

Cuando nos movemos del ámbito de la experiencia humana a la experiencia animal, desde una perspectiva externa se puede aplicar el mismo tipo de análisis. El caso extremo resulta ilustrativo. Hace bastante tiempo que los biólogos saben que los invertebrados tienen un repertorio conductual más bien limitado, Por ejemplo, la cucaracha sólo posee unas pocas formas fundamentales de locomoción: detenerse, caminar lentamente, caminar rápidamente y correr. Sin embargo, este repertorio conductual básico permite que estos animales funcionen adecuadamente en cualquier ambiente posible conocido en el planeta, ya sea natural o artificial. Surge una interrogante clave para el biólogo: ¿cómo decide el animal la actividad motriz que debe emprender en una

[5] Me estoy refiriendo específicamente a *Being and Time* de Martin Heidegger, trad,. de John Mac-quarrie y Edward Robinson (Nueva York: Harper, 1929) y *Phenomenology of Perception* de Merleau-Ponty, trad. de Colin Smith (Nueva York: humanities Press, 1962)

situación dada? ¿Cómo opera su selección conductual para que la acción sea efectiva? ¿Cómo posee el animal el sentido común para evaluar una situación dada e interpretar que exige que corra en vez de caminar lentamente? En estos dos casos extremos —la experiencia humana en los momentos de quiebre, y la conducta animal simple en momentos de transición conductual— nos topamos, en forma muy diferente, claro está, con un problema común. En cada uno de estos momentos de quiebre, el modo en que volverá a constituirse el agente cognitivo no es materia de decisión externa ni tampoco de planificación previa. Más bien, se trata de la *emergencia del sentido común*, de la configuración autónoma de una postura adecuada. Una vez que ha sido seleccionada una postura conductual o que se produce un micromundo, podemos analizar con mayor claridad su forma de operación y su estrategia óptima. De hecho, la clave de la autonomía es que un sistema viviente transcurre hasta el momento siguiente mediante una acción efectiva que surge de sus propios recursos. Y estas rupturas, bisagras que articulan los micromundos, están en el origen del carácter autónomo y creativo de la cognición en los seres vivos. Es necesario entonces examinar a microescala este sentido común, ya que lo concreto se actualiza durante los quiebres.

El conocimiento como enacción. Es necesario explicar lo que quiero decir con la palabra "encarnado", y para ello, hay dos puntos que quisiera enfatizar. En primer lugar, la cognición depende de los tipos de experiencia que provienen del hecho de tener un cuerpo con varias habilidades sensorimotrices. En segundo lugar, estas habilidades sensorimotrices individuales se alojan, a su vez, en un *contexto* biológico y cultural más amplio. Ya abordamos estos dos puntos cuando hablamos de las rupturas y del sentido común, pero deseo explorar más a fondo su especificidad corpórea y enfatizar que los procesos sensorimotrices, la percepción y la acción, son fundamentalmente inseparables en la cognición vivida, y *no* están simplemente ligados en forma contingente.

Al adoptar lo que he denominado un "enfoque enactivo de la cognición"[6], habrá que destacar dos principios fundamentales. Primero, la percepción es acción perceptualmente guiada; y segundo, las estructuras cognitivas surgen de los esquemas sensorimotrices recurrentes que permiten que la acción sea guiada perceptualmente. (Estos puntos quedarán más claros a medida que avanzo).

Comenzaré por el concepto de acción perceptualmente guiada. Para la tradición computacional dominante, el punto de partida para comprender la percepción es típicamente abstracto: el problema del procesamiento de la información para recuperar propiedades del mundo pre-determinadas. En contraste, el punto de partida para el enfoque enactivo es el estudio de cómo el que percibe guía sus acciones en situaciones locales. Dado que estas situaciones locales cambian constantemente como resultado de la actividad del que percibe, el punto de referencia para comprender la percepción ya no es un mundo pre-dado, independiente del que percibe, sino más bien la estructura sensorimotriz del agente cognitivo, el modo en que el sistema nervioso une superficies sensoriales y motrices. Más que un mundo pre-determinado, es esta estructura —la forma en que el sujeto que percibe está *encarnado*- la que determina los modos de acción posibles del que percibe y cómo éste puede ser moldeado por acontecimientos ambientales. De este modo, la preocupación esencial de un enfoque enactivo de la percepción no es determinar cómo se recupera un mundo independiente del que percibe; se trata, más bien, de determinar los principios comunes o las vinculaciones entre el sistema sensorial y el motor que explican cómo la acción puede ser *perceptualmente* guiada en un mundo *dependiente del que percibe*.

[6] Francisco Varela, *Connaître: Les Sciences cognitives* (París: Seuil, 1989); Varela, "Organism: A Meshwork of selfless selves," en Alfred Tauber, ed., Organism and the Origin of Self (Dordrecht and Uitgeverij: Reidel Kluwer, 1991); Varela, Evan Thompson & Eleanor Rosh, Th*e Embodied Mind: Cognitive Science and Human Experience* (Cambridge, Mass.: MIT Press, 1991); y Thompson, Alden Palacios and Varela, "Ways of coloring: Comparative Color Vision as a Case Study in the Foundations of Cognitive science," *Behavioral Brain Sciences* 16.1 (en prensa).

Esta preocupación fundamental de la posición enactiva se opone al planteamiento de que la percepción es fundamentalmente un registro de información ambiental existente para reconstruir fidedignamente un aspecto del mundo físico. La realidad no es algo dado: es dependiente del que percibe, no porque éste "construye" el mundo en forma antojadiza, sino porque lo que cuenta como mundo relevante es inseparable de la estructura del que percibe.

De hecho, esta perspectiva de la percepción constituye uno de los hallazgos fundamentales del análisis fenomenológico de Maurice Merleau-Ponty en sus trabajos iniciales. Vale la pena que citemos uno de sus pasajes más visionarios:

> No podemos comparar al organismo con un teclado sobre el que actuarían los estímulos externos y en que su forma correcta aparecería delineada por la sencilla razón que el organismo contribuye a constituir esa forma... "Las propiedades del objeto y las intenciones del sujeto... no sólo se entremezclan, sino que constituyen un nuevo "todo". Cuando el ojo y el oído siguen la trayectoria de un animal en vuelo, es imposible determinar "cuál fue primero" en el intercambio de estímulo y respuesta. Dado que todos los movimientos del organismo están siempre condicionados por influencias externas, si quisiéramos, podríamos considerar que la conducta es un efecto del medio. Pero, del mismo modo, ya que lo que ha posibilitado los estímulos que recibe el organismo han sido sus movimientos precedentes, que han culminado en la exposición del órgano receptor a las influencias externas, también podríamos decir que la *conducta es la primera causa de toda estimulación*.
>
> De este modo, la forma de la exitación es *creada* por el propio organismo, por su propia forma de ofrecerse a las acciones provenientes del medio. Sin duda, para poder susbsistir, debe hayar ciertos agentes físicos y químicos. Pero es el propio organismo —de acuerdo a la naturaleza de sus receptores, el umbral de sus centros nerviosos y los movimientos de sus órganos— *que elige los estímulos del mundo físico ante los cuales reaccionará*. "El ambiente *(Umwelt)* emerge del mundo mediante la actualización o el ser del organismo —ya que el organismo sólo puede existir si encuentra en el mundo un medio ambiente apropiado". Este sería un teclado que se mueve a sí mismo en forma tal que ofrece —y de acuerdo a ritmos variables—

esta o esta otra tecla a lo que en sí no es más que la acción monótona de un martillo externo[7].

Según este enfoque, la percepción no está simplemente alojada en, y constreñida por el mundo circundante, sino que contribuye a la *enacción* de este mundo circundante. Como señala Merleau-Ponty, el organismo se inicia en el ambiente y, a la vez, es moldeado por él, reconociendo claramente que el organismo y el ambiente están estrechamente unidos en una especificación y selección recíprocas —un punto que debiéramos recordar constantemente, ya que está en abierta contradicción con la perspectiva imperante derivada de la tradición cartesiana.

El estudio de Held y Hein es una ilustración clásica de la acción guiada perceptualmente. Estos investigadores criaron gatitos en la oscuridad y los expusieron a la luz sólo en condiciones controladas[8]. A un primer grupo de animales se le permitió desplazarse normalmente, pero estaban enganchados a un carro sencillo y a un canasto que contenía al segundo grupo de animales. Por lo tanto, ambos grupos compartieron la misma experiencia visual, pero el segundo grupo era completamente pasivo. Cuando los animales fueron puestos en libertad unas cuantas semanas después de este tratamiento, el primer grupo de gatitos se comportó normalmente, pero los que habían sido transportados se comportaban como si estuviesen ciegos. Chocaban con los objetos y perdían el equilibrio. Este estudio respalda el planteamiento de la percepción como *acción* según el cual los objetos no son vistos mediante la extracción visual de características, sino más bien por la acción guiada visualmente. Se han obtenido resultados similares en muchas otras circunstancias y han sido estudiados incluso a nivel de una sola célula.

---

7 Maurice Merleau-Ponty, *The structure of Behavior*, trad. de Alden Fisher (Boston: Beacon, 1963), p. 13 (énfasis nuestro). La primera cita interna es de V.F. von Weizsäcker, "Reflexgesetze," en Bethe, ed., *Handbuch der normalen und pathologischen physiologie*, pp. 38-39: la segunda es de K. Goldstein, *The Organism* (Boston: Beacon, 1963).

8 Richard Helm y Alan Hein, "Adaptation of Disarranged Hand-Eye Coordination Contingent upon Re-afferent Stimulation," *Perceptual and Motor Skills* 8 (1958), pp. 87-90.

A menos que el lector opine que este ejemplo sirve para los gatos, pero que no tiene relación con la experiencia humana, tomemos otro caso. En 1962, Paul Bach y Rita diseñaron una cámara de video para ciegos que estimula múltiples puntos de la piel mediante vibración eléctricamente activada[9] . Utilizando esa técnica, se logró que las imágenes formadas con la cámara correspondieran a patrones de estimulación de la piel, sustituyendo de este modo la pérdida visual. Los patrones proyectados sobre la piel no tienen contenido "visual", a menos que el individuo sea conductualmente activo manipulando la cámara de video con movimientos de la cabeza, la mano o del cuerpo. Cuando la persona no vidente se comporta activamente de ese modo, después de unas cuantas horas de experiencia ocurre algo sorprendente: la persona ya no interpreta las sensaciones de la piel como relacionadas con el cuerpo, sino más bien como imágenes proyectadas en el espacio que está siendo explorado por la "mirada" corporalmente dirigida de la cámara de video. Por lo tanto, para poder tener la experiencia de "verdaderos objetos ahí afuera", la persona debe dirigir activamente la cámara (utilizando la cabeza o la mano).

La fina estructura del presente. Después de haber situado la emergencia de lo concreto en el contexto de la perspectiva de la cognición como acción, donde adquiere todo su sentido, podemos volver a la pregunta inicial. ¿Cómo surge uno de los tantos micromundos potenciales de la multitud de agentes cognitivos y subredes? La respuesta que deseo proponer es que al interior de la brecha que ocurre durante un quiebre, existe una rica dinámica que involucra subentidades y agentes concurrentes. Este rápido diálogo, no accesible a la introspección, ha sido recientemente revelado en estudios del cerebro.

Algunos de los aspectos clave de esta idea fueron introducidos por Walter Freeman, quien, a lo largo de varios años de investigación, pudo colocar una serie de electrodos en el bulbo olfatorio de un conejo para medir pequeñas porciones de actividad global mientras

---

9 Paul Bach y Rita, ***Brain Mechanisms in Sensory Substitution*** (New York: Academic Press, 1972).

el animal se comportaba libremente[10]. Descubrió que no existe un patrón claro de actividad global en el bulbo, a menos que se exponga varias veces al animal a un olor específico. Además, descubrió por primera vez que estos patrones emergentes de actividad se crean a partir de un trasfondo de actividad incoherente o caótica mediante oscilaciones rápidas (por ejemplo, con períodos de alrededor de 5-10 milisegundos) hasta que el córtex se acomoda en un patrón eléctrico global, que dura hasta el final de la actividad olfatoria y luego se disuelve en el trasfondo caótico[11]. Las oscilaciones son, entonces, el medio para *unir selectivamente* un conjunto de neuronas en un agregado transitorio que constituye el sustrato de la percepción del olor en ese preciso instante. Bajo esta perspectiva, el olor aparece, no ya como un mapeo de características externas, sino como una forma creativa de materializar el significado sobre la base de la historia encarnada del animal. Aquí, lo más pertinente es que esta materialización ocurre en el gozne entre un momento conductual y el siguiente, a través de oscilaciones rápidas entre poblaciones neuronales que pueden hacer surgir patrones coherentes.

Cada vez hay más evidencia sobre este tipo de resonancia rápida que une transitoriamente conjuntos de neuronas durante un percepto. Esta ha sido observada, por ejemplo, en la corteza visual de los gatos y monos durante la estimulación visual[12]; también se la ha encontrado en estructuras neurales radicalmente diferentes, tales como el cerebro del pájaro[13] e incluso en los ganglios del invertebrado *Hermissenda*[14].

---

10 Walter Freeman, *Mass Action in the Nervous System* (Nueva York: Academic Press, 1975).

11 Walter Freeman y Christine Skada, "Spatial EEG Patterns, Nonlinear Dynamics, and Perception: The Neo-Sherringtonian View," *Brain Research Reviews* 10 (1985), pp. 147-75.

12 Charles Gray y Wolf Singer, "Stimulus-Specific Neuronal Oscillations in Orientation Columns in Cat Visual Cortex," *Proceedings of the National Academy of Sciences of the USA* 86 (1989), pp. 1698-702.

13 Serge Neuenschwander y Francisco Varela, "Sensori-triggered and Spontaneous Oscillations in the Avian Brain," *Society Neuroscience Abstracts* 16 (1990).

14 Alan Gelperin y David Tank, "Odour-Modulated Collective Network Oscillations of Olfactory Interneurons in a Terrestrial Mollusc, " *Nature* 345 (1990), pp. 437-40. Para un análisis reciente, véase Steven Bressler, "The Gamma Wave: A Cortical Information Carrier," *Trends in Neuroscience* 13 (1990), pp. 161-62.

Esta universalidad es importante ya que apunta a la naturaleza fundamental de la vinculación producto de la resonancia como un mecanismo que permite el surgimiento de los acoplamientos sensorimotores. Si se hubiese tratado de un proceso muy específico de la especie, típico por ejemplo de la corteza de los mamíferos, sería mucho menos interesante como hipótesis de trabajo.

Es importante destacar que esta resonancia rápida no está vinculada simplemente a un gatillamiento sensorial. Las oscilaciones aparecen y desaparecen rápidamente y en forma bastante espontánea en varios lugares del cerebro. Esto sugiere que esta rápida dinámica involucra a la totalidad de las redes que dan origen a toda disposición a la acción en el siguiente momento. No involucra solamente una interpretación sensorial y una acción motriz, sino también toda la gama de expectativas cognitivas y la tonalidad emocional, que son esenciales para la conformación de un microuniverso. Entre rupturas, estas oscilaciones constituyen los síntomas de una muy rápida cooperación y competencia recíprocas entre distintos agentes que son activados por la situación en curso, rivalizando entre sí respecto de distintos modos de interpretación para un marco cognitivo coherente y una disposición a la acción. sobre la base de esta rápida dinámica, al igual que en un proceso evolutivo, un conjunto neuronal (una subred cognitiva) finalmente se vuelve más predominante y *se transforma en el modo de conducta para el siguiente momento cognitivo.* Cuando digo "se vuelve predominante" no quiero decir que se trata de un proceso de optimización: se parece más bien a una bifurcación o a la ruptura de la simetría en una dinámica caótica. De ahí se desprende que semejante cuna de la acción autónoma queda excluida para siempre de la experiencia vivida, ya que, por definición, sólo podemos habitar una microidentidad cuando está presente y no cuando está siendo gestada. En otras palabras, en el quiebre que se produce antes que surja el siguiente micromundo, hay toda una gama disponible de posibilidades hasta que, de las exigencias de la situación y de la recurrencia de la historia, se selecciona una sola. Esta rápida dinámica es el correlato neural de la constitución autonóma de un agente cognitivo incorporado en un momento dado de su vida.

De la fina estructura temporal a la acción cognitiva. Como ya señalamos, la rápida resonancia de la reciprocidad de un agente proporciona el trasfondo para la emergencia de un micromundo. La evidencia sugiere que este acoplamiento sensorimotriz está vinculado con otros tipos de ejecución cognitiva típicamente humanos, lo que significa que los niveles cognitivos más "altos" surgen de los acontecimientos de "bajo" nivel como sentir y actuar, permitiendo que la acción sea guiada perceptualmente.

De hecho, esta idea básica constituye el núcleo del programa piagetiano[15] y ha sido defendida en varios trabajos recientes de George Lakoff y Mark Johnson[16]. Presentaré la idea de estructuras cognitivas encarnadas con especial referencia a su trabajo. Una vez más, debemos salir de lo abstracto y enfatizar un enfoque experiencial de la cognición. Como afirma Lakoff, el planteamiento central de este enfoque es que las estructuras conceptuales significativas surgen de dos fuentes: de la naturaleza estructurada de la experiencia corporal, y de nuestra capacidad para proyectar imaginativamente desde algunos aspectos bien estructurados de la experiencia corporal e interaccional a estructuras conceptuales.

El pensamiento racional y abstracto es una aplicación de procesos cognitivos muy generales —focalización, scanning, superposición, inversión figura-fondo, etc.— a las estructuras conceptuales[17]. La idea fundamental es que las estructuras encarnadas (sensorimotrices) son el meollo de la experiencia, y que las estructuras experienciales "motivan" la comprensión conceptual y el pensamiento racional. Ya destaqué que la percepción y la acción están encarnadas en los procesos sensorimotores (que son autoorganizantes): por lo tanto, es natural postular que las estructuras cognitivas *surgen* de esquemas recurrentes

---

15 Jean Piaget, *Biologie et Connaissance* (París: Gallimard, 1969).

16 George Lakoff, *Women, Fire and Dangerous Things* (Chicago: University of Chicago Press, 1983); y Mark Johnson, *The Body in the Mind* (Chicago: University of Chicago Press, 1989).

17 George Lakoff, "Cognitive Semantics" en Umberto Eco et al., eds., *Meaning and Mental Representations* (Bloomington: Indiana University Press, 1988), p. 121 proporciona un análisis conciso del enfoque experiencialista de Lakoff y Johnson.

de actividad sensorimotriz. En cualquier caso, no es que la experiencia determine estrictamente las estructuras conceptuales y los modos de pensamiento (tal como lo plantea Lakoff): más bien el punto es que la experiencia hace posible y a la vez constriñe la comprensión conceptual a través de los diferentes dominios cognitivos.

Lakoff y Johnson proporcionan numerosos ejemplos de estructuras cognitivas generadas a partir de procesos experienciales. Revisar todos esos ejemplos nos alejaría mucho del tema central. Permítanme, sin embargo, examinar brevemente uno de los tipos más significativos: las categorías de nivel básico. Piensen en la mayoría de las cosas de tamaño intermedio con las que interactuamos continuamente: mesas, sillas, perros, gatos, tenedores, cuchillos, tazas, etc. Estas cosas pertenecen a un nivel de categorización intermedia que se encuentra entre los niveles más bajos (subordinado) y los niveles más altos (superordinado). Si tomamos, por ejemplo, una silla en el nivel más bajo podría pertenecer a la categoría "mecedora", mientras que en el nivel más alto pertenece a la categoría "mueble". Eleanor Rosch *et al* demostraron que este nivel de categorización intermedia (mesa, silla, etc.) psicológicamente es el nivel más fundamental o *básico*, por las siguientes razones, entre otras: en primer lugar, el nivel básico es el nivel más general en que los miembros de la categoría poseen *formas percibidas* similares; segundo, es el nivel más general en que una persona utiliza acciones motrices similares para interactuar con miembros de la categoría; tercero, es el nivel en que los grupos de atributos correlacionados son más *aparentes*[18]. Por lo tanto, pareciera que lo que determina que una categoría pertenezca al nivel básico no depende de cómo las cosas están ordenadas en un mundo pre-establecido, sino más bien de la estructura sensorimotriz de nuestros cuerpos y de los tipos de interacciones perceptualmente guiadas que esta estructura permite. Las categorías de nivel básico son a la vez experienciales y encarnadas. Se puede esgrimir un argumento similar para los esquemas-imágenes que surgen de formas básicas de actividades e interacciones sensorimotrices.

---

[18] Eleanor Rosch, Carolyn Mervis, Wayne Gray, David Johnson y Penny Boyes-Braem, "Basic Objects in Natural Categories," *Cognitive Psychology* 8 (1976), pp. 382-439.

3. Conclusión

He argumentado que la percepción no consiste en recuperar un mundo pre-dado, sino que consiste más bien en una acción perceptualmente guiada en un mundo que es inseparable de nuestras capacidades sensorimotrices. Las estructuras cognitivas surgen de pautas recurrentes de acción perceptualmente guiadas. Para resumir entonces, la cognición no consiste en representaciones, sino en *acciones encarnadas*. Correlativamente, podemos decir que el mundo que conocemos no es un mundo pre-establecido; más bien es un mundo *enactuado* a través de nuestra historia de acoplamiento estructural, y los goznes temporales que articulan la emergencia de diferentes modos de acción están enraizados en rápidas dinámicas no-cognitivas que activan diferentes micromundos. Estos goznes constituyen, a la vez, la fuente del sentido común y de la creatividad en cognición.

Es así como la propia búsqueda contemporánea de la ciencia cognitiva para comprender el comprender apunta en una dirección que podemos considerar post-cartesiana. En primer lugar, cada vez hay más evidencia de que el conocimiento se construye desde pequeños dominios; es decir, desde los micromundos y las microidentidades. Estas modalidades básicas de disposición-a-la-mano son variables a través del reino animal. Pero lo que todos los seres vivos cognitivos parecen tener en común es conocimiento, que es siempre un *know-how* constituido sobre la base de lo concreto; lo que denominamos lo "general" y lo "abstracto" son agregados de la disposición a la acción. Segundo, estos micromundos no son coherentes ni están integrados en una totalidad que regula a las instancias más pequeñas: es justamente esta falta de orden lo que permite la emergencia de un momento cognitivo de acuerdo a la historia y la constitución del organismo. El meollo de esta autonomía, la rapidez con que el agente selecciona una conducta, es totalmente inaccesible al sistema cognitivo en cuestión. De este modo, lo que tradicionalmente ha sido denominado "irracional" e "inconsciente" no está en contradicción con lo que aparece como racional y volitivo: al contrario, es su verdadero fundamento.

II

MENTE

2. Conciencia

CUATRO PAUTAS PARA EL FUTURO DE LAS CIENCIAS COGNITIVAS

Escribí este texto para la conferencia inaugural de la IIª Media Bienal, en Hanover. Para poder dirigirme a un público tan diverso y sin bases necesariamente científicas, escogí un tono de reflexión sobre el futuro, propio de una bienal. El resultado es también una reflexión sobre la transición entre pasado y presente, sobre el futuro del conocimiento, lo que constituye un buen preludio a los temas más arduos de esta última Sección, dedicada a la conciencia.

Trabajo con el cerebro, estudiando la neurociencia y las raíces biológicas del conocimiento, de modo que, cuando recibí la invitación para esta conferencia, me pregunté: ¿Qué sería lo más útil para esta discusión? ¿Cuáles son los conceptos que podrían llevarse de esta conferencia y que, sin ser ustedes científicos, pudiesen servirles para su propia labor? En definitiva, lo que me gustaría hacer en la siguiente media hora es un recorrido de lo que considero los principios-guía, los puntos-clave de la evolución de este ámbito de la ciencia consagrado al estudio del cerebro y de la cognición —hacia adónde va: cuáles son nuestros principales logros y cuáles son las cosas que todavía nos quedan por hacer. No voy a justificar todo esto, porque obviamente no hay tiempo

---

(Varela, F. (2000), Four batons for the future of cognitive science, in *Envisioning Knowledge*, B.Wiens (Ed.), Dumont, Cologne, 1999). Este trabajo es una versión condensada del artículo de F. Varela (1999), Steps to a Science of Interbeing: Unfolding the Dharma implicit in modern cognitive science, en S. Bachelor, G, Claxton y G. Watson (eds.), *The Psychology of Awakening*, Rider/Random House, Nueva York.)

para ello, sino que simplemente se los voy a lanzar como un listado de puntos-clave. Representan cuatro pautas, puntos-clave para el futuro del conocimiento. Este es el ejercicio que hoy me propongo. ¿Qué es la ciencia cognitiva para *ustedes*? Les voy a presentar estos puntos-clave en forma de lemas, o, para ser más preciso, esquemas fundamentales, fruto de unos cincuenta años de investigación, y a los que considero resultados establecidos. Después de muchos años de trabajo en los laboratorios de ciencia cognitiva, se podría decir que les estoy presentando mensajes envasados. Estas no son interpretaciones muy personales, sino un intento por extraer aquellos mensajes fundamentales que pueden ser seriamente defendidos. Estoy tratando de hacer una lectura congruente con los resultados empíricos.

1. El punto-clave de la encarnación

El primer punto es lo que denomino el punto clave de la encarnación. Esto en contraste con la perspectiva imperante hoy en día basada en la metáfora computacional, en que la mente es considerada como el *software*, y el cerebro y el cuerpo como el *hardware*. Lo que aquí denomino mente es cualquier fenómeno relacionado con la mentalidad, con la cognición, y en último término con la experiencia. Uno de los más importantes avances en ciencia en los últimos años es la convicción de que no podemos tener nada que se asemeje a una mente o a una capacidad mental sin que esté totalmente encarnada o inscrita corporalmente, envuelta en el mundo. Surge como una evidencia inmediata, inextricablemente ligada a un cuerpo que es activo, que se mueve y que interactúa con el mundo.

Puede que esto les parezca obvio, pero en el mundo de la investigación han prevalecido otras ideas, especialmente la metáfora computacional a la que antes me referí. Es necesario romper con esa tradición ya que hay toda una serie de argumentos para afirmar que la mente no es un programa, un *software*, una manipulación de símbolos basada en determinadas reglas. Para resumir brevemente, aquí va mi primer lema: *la mente no está en la cabeza*. Esta idea surge como consecuencia del redescubrimiento de la importancia del estar encarnado.

Ahora bien, ¿por qué es esto así? Responde a una lógica muy precisa: una vez que hayan comprendido que para que exista una mente tiene que haber manipulación e interacción activa con el mundo, entonces tenemos un fenómeno incorporado y activo, y cualquier cosa que denominemos un objeto, una cosa en el mundo, las sillas y mesas, las personas y las caras y todo lo demás, depende totalmente de esta constante manipulación sensorimotriz. No podemos captar al objeto como si simplemente estuviera "ahí afuera" en forma independiente. El objeto surge como fruto de nuestra actividad, por lo tanto, tanto el objeto como la persona están co-emergiendo, co-surgiendo.

Voy a presentarles un ejemplo para que vean que no se trata simplemente de algo metafórico: hace algún tiempo, Held y Hein llevaron a cabo un experimento clásico con dos gatitos, ciegos al nacer, en dos canastas. Cada gatito fue colocado dentro de una canasta y cada día eran paseados durante algunas horas dentro de ésta; es decir, ambos gatitos fueron expuestos al mismo ambiente. A uno de los gatitos se le permitió que mantuviera las patas fuera de la canasta y que caminara, al otro se lo mantuvo arropado dentro de ésta. Dos meses después los gatitos fueron puestos en libertad. El gatito al que se le había permitido caminar se comportó como un gato normal. El otro no reconocía los objetos, se caía por las escaleras y chocaba contra las sillas. Prácticamente, se comportaba como si estuviera ciego, aunque sus ojos estaban intactos. ¡La conclusión que *no* hay que sacar es que los gatos ven con los pies! La conclusión que hay que sacar es que el espacio surge como producto del movimiento. Esta es una constatación absolutamente extraordinaria: el espacio, esta cosa frente a nosotros que parece absolutamente objetiva, el pilar de la objetividad en física, es totalmente inseparable del hecho que tenemos que manipularlo a través de una conducta sensorimotriz. Hay un sinnúmero de ejemplos que no puedo describir aquí que permiten desarrollar este mismo tipo de argumento. En mi libro *De cuerpo Presente*[1] hay una descripción detallada de cómo esta noción

[1] F. Varela, E. Thompson y E. Rosch (1991), *The Embodied Mind: Cognitive science and human experience*, MIT Press, Cambridge.

se puede aplicar a la percepción y a la calidad del color, nuevamente una propiedad totalmente *co-emergente.*

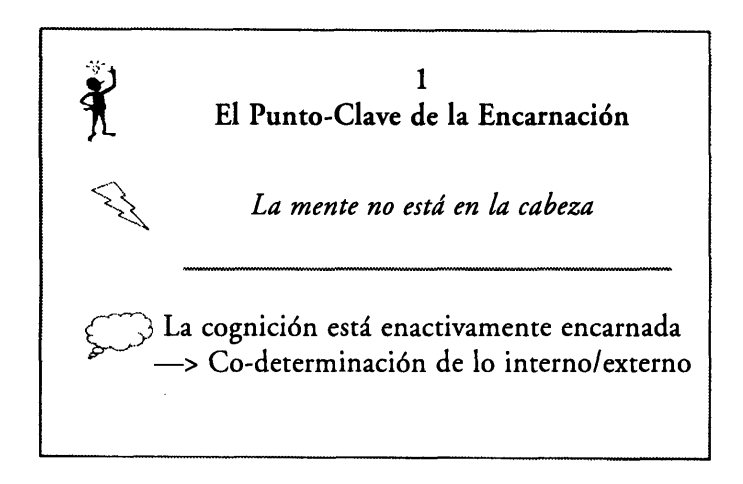
Transformemos este lema "*La mente no está en la cabeza*" en una lógica más estructurada: La cognición está enactivamente encarnada. "Enactiva" es una etiqueta que utilizo aquí en su sentido literal ya que la cognición es algo que producimos por el acto de manipular, por medio de una manipulación activa: es el principio fundacional de lo que es la mente. Como traté de mostrar anteriormente, esto implica una profunda co-implicación, una co-determinación entre lo que parece estar afuera y lo que parece estar adentro. En otras palabras, el mundo ahí afuera y lo que hago para estar en ese mundo son inseparables. El proceso los vuelve totalmente interdependientes, como vimos en el ejemplo de los gatitos.

Esta perspectiva de la mente como enactivamente encarnada tiene dos consecuencias ya que, si la mente no está en la cabeza, ¿dónde diablos está? Este es precisamente el punto: es en este no-lugar de la co-determinación entre lo interno y lo externo, luego no podemos decir que está afuera o adentro. La otra consecuencia que se deriva de esto y que ha sido menos enfatizada, es que la mente es inseparable del organismo como *un todo.* Tendemos a creer que la mente está en el cerebro, en la cabeza, pero el hecho es que el ambiente también incluye al resto del organismo: incluye el hecho de que el cerebro está íntimamente conectado con todos los músculos, con el esqueleto, los intestinos, y el sistema inmunitario, los flujos hormonales y así sucesivamente. Hace de todo el conjunto una unidad sumamente apretada. En otras palabras, el organismo como una red de elementos totalmente co-determinados determina que nuestra mente sea, literalmente, inseparable, no sólo del ambiente externo, sino también de aquello que Claude Bernard denominó el *milieu intérieur*, el hecho de que no sólo estamos dotados de un cerebro sino de todo un cuerpo.

Si ustedes provienen de esa tradición de la filosofía de la mente según la cual la mente es algo que ocurre dentro de la cabeza, puede que esto les parezca bastante sorprendente. Por ejemplo, en el pasado, los filósofos se entretuvieron hablando de cerebros en un recipiente,

un cerebro en un tubo de ensayo provisto de alambritos. Es divertido, la comunidad filosófica anglo-americana ha pasado muchas horas discutiendo este tipo de cosa, pero cuando vemos el estado del arte de la investigación, todo el argumento parece extraño porque con una mente en un tubo de ensayo no puede haber mente. Lo que tendríamos es una actividad neural completamente incoherente, porque no podría tener la funcionalidad de lo que realmente hace, el manejo y la interacción constante con el cuerpo y con el ambiente que le da sentido.

En resumen, ese es el primer punto de la mente encarnada, la co-determinación de lo interno y lo externo. Y no hay que olvidar que esto se refiere tanto al ambiente externo como al propio cuerpo.



2. El punto-clave de la emergencia

El segundo lema es: *la mente ni existe ni no existe.* Esto es lo que denomino el punto-clave de la emergencia. Esta es una noción difícil, porque la noción de emergencia es fundamental y por lo general ha sido incomprendida. Utilizo el término emergencia en un sentido más bien técnico. Cuanto más observamos el funcionamiento del cerebro, o más bien, el funcionamiento del proceso de cognición, cada vez hay más evidencia de que estamos tratando con componentes muy individuales, neuronas o grupos de neuronas,

o poblaciones de neuronas. El trabajo del neurocientífico es investigar estas células y tratar de comprender su funcionamiento, la impresionante riqueza del cerebro y la extraordinaria complejidad de millones y millones de intrincadas conexiones. Podemos referirnos a estos elementos locales en interacción como las reglas locales; estas reglas locales e interacciones locales no son como el traspaso de información en un computador —el envío de mensajes de un lado para otro en forma sintáctica o programática. Estas interacciones ocurren en un tiempo real en forma muy rápida, dinámica y simultánea.

Lo que se desprende de lo anterior es algo que todavía no termina de asombrarme, y es que a partir de este elemento local surge un proceso *global*, un estado global o nivel global, que ni es independiente de estas interacciones locales ni reducible a ellas. Se trata de la emergencia de un nivel global que surge a partir de las reglas locales y que tiene un status ontológico diferente, porque trae consigo la creación de un individuo, o de una unidad cognitiva. Entonces, cuando ven a un animal en movimiento, o me ven a mí dirigiéndoles la palabra, me comporto como una unidad coherente, no como una mera yuxtaposición de movimiento, voz, mirada, y postura. Soy una unidad integrada, más o menos armónica a la que denomino "mí mismo" o "mi" mente, y ustedes interactúan conmigo a ese nivel: "Hola Francisco". Esa interacción está ocurriendo a nivel de la individualidad, que es lo global, lo emergente. Sin embargo, sabemos que lo global es al mismo tiempo causa y consecuencia de las acciones locales que ocurren todo el tiempo en mi cuerpo.

Ahora bien, no tengo acceso a la frecuencia de descarga de mi neurona 223 en la corteza visual, porque esto sería como ir en contra de la ley según la cual lo global es ontológicamente distinto a lo local. La noción de emergencia, un principio general que ha impregnado todas las ciencias y no sólo a la neurociencia durante los últimos veinte años, revela la fundamental importancia de concebir un nuevo modo o tipo de existencia, una nueva forma de caracterizar lo que es una cosa. Es un modo de existencia del cual no puede decirse que no existe ("Francisco no existe"). Yo cuento para algo, ustedes están leyendo lo que escribí. Al mismo tiempo, ¿cuál es la naturaleza de mi

existencia? No asumimos que hay algo sustancial o una cualidad especial que está alojada en alguna parte de mi cerebro que hace que Francisco sea Francisco. En realidad, este *yo* cognitivo es fruto de sus acoplamientos dinámicos, los que incluyen a todos y cada uno de los componentes locales, pero al mismo tiempo, no puede reducirse a ninguna interacción en particular. Por lo tanto, es como decir que está y no está ahí.

Es como cuando decimos "Inglaterra". En un sentido Inglaterra existe, tiene vínculos comerciales, tratados de paz, y hace todo tipo de cosas como nación. Al mismo tiempo. ¿dónde está? Inglaterra está en la pauta de interacciones de la gente que vive ahí. Luego, ni existe ni tampoco no existe, esto es, sólo existe como un patrón en flujo. Claramente, tiene un *modo* de existencia que no es del tipo que hemos heredado de la teología o la física, concebida como sustancialidad o materialidad, una cualidad que podemos ubicar o localizar de alguna u otra forma. Sin embargo, las entidades emergentes son la base de entidades complejas típicas de los ámbitos de la vida y de la mente, sumamente efectivas como modo de acción y modo de presencia en el mundo.

Por lo tanto, la cognición no sólo está enactivamente encarnada, sino que también es *enactivamente emergente*, en ese sentido técnico que traté de esbozar. Puede que algunas personas se refieran a este fenómeno bajo diferentes nombres: autoorganización, complejidad, o dinámica no lineal. El principio de base es el mismo: el tránsito de lo local a lo global. Es una codeterminación entre elementos neurales y un sujeto cognitivo global. El sujeto cognitivo global pertenece a ese nivel emergente y a ese modo de existencia.

Ahora bien, el principio de la emergencia normalmente se interpreta en forma más bien reduccionista. Lo que quiero decir es que muchas personas aceptan que el *yo* es una propiedad emergente que surge de una base neural/corporal. Sin embargo, la afirmación inversa generalmente se pasa por alto. Esto es importante ya que, si los componentes y circuitos neurales funcionan como agentes locales, capaces de dar origen a un *yo* en forma emergente, de ello se deriva que este nivel global, el *yo*, tiene acciones eficaces directas

sobre los componentes locales. Es una calle de dos direcciones: los componentes locales dan origen a esta mente emergente, pero, vice versa, la mente emergente constriñe, afecta directamente a los componentes locales.

Desde este punto de vista, el enigma de los fenómenos psicosomáticos es un falso problema. ¿Por qué habría de sorprendernos que un estado global correspondiente a un estado cognitivo o a una actitud o a un estado mental puedan tener un efecto directo en principios locales muy, pero muy diminutos? Y si creen que ésta no es más que una aspiración, permítanme un ejemplo. Hemos estado trabajando con pacientes epilépticos a los que se les implantan electrodos en el cerebro unos días antes de ser operados. De este modo, tenemos acceso a señales eléctricas muy precisas del cerebro de un humano en estado de vigilia, y también podemos estudiar el momento que precede a la crisis, así como también predecir su ocurrencia con algunos minutos de anticipación[2]. Este es un un buen ejemplo de cómo las propiedades locales (las corrientes locales) conducen a un estado global (la crisis). Pero también comprobamos lo contrario: cuando un paciente está activamente concentrado en una actividad cognitiva (como reconocer una forma visual) se detectan cambios en la dinámica epiléptica. Esto significa que el efecto de un estado global tiene efectos sobre la actividad eléctrica local en forma my precisa[3]. Hablemos entonces sobre la integración mente-cuerpo. Por alguna extraña razón, tenemos esta tradición occidental, esta percepción de que la materia puede ser el soporte de la mente, pero no que la mente pueda tener un efecto directo sobre la materia. Bueno, esta es una noción demostrablemente equivocada, y esto es lo que el concepto de emergencia nos permite apreciar, si lo entendemos correctamente como una calle de dos direcciones y no tan sólo de una dirección.

---

[2] J. Martinerie, C. Adam, M. Le van Quyen, M. Baulac, B. Renault y F. J. Varela (1998), Epileptic crisis can be anticipated by non-linear analysis, *Nature Medicine* 4:1173-1176.

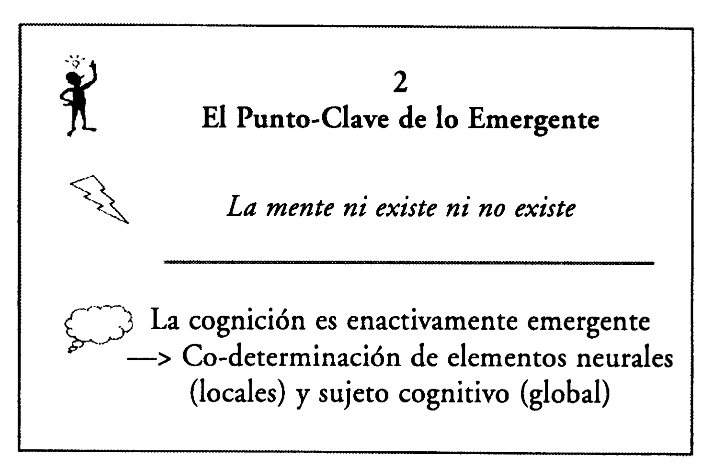
[3] Le Van Quyen, M. J. Martinerie, C. Adam, J-Ph. Lachaux, M. Baulac, B. Renault, y F. Varela (1997) Temporal patterns in human epileptic activity are modulated by perceptual discriminations, *Neuroreport* 8:1703-1710.

Por lo tanto, la cognición es enactivamente emergente y es la co-determinación entre elementos locales y el sujeto cognitivo global. Hay dos corolarios que se desprenden de este punto clave. En primer lugar, si juntamos el primer punto clave y el segundo, es decir, los conceptos de encarnación y de emergencia, la mente es fundamentalmente asunto de imaginación y de fantasía. En otras palabras, es la actividad interna de estas propiedades emergentes, junto con el hecho de que hay un acoplamiento continuo, lo que constituye el aspecto central de lo que es la mente. La mente no es la representación de un determinado estado de cosas; la mente es la producción constante de esta realidad coherente que constituye un mundo, un modo coherente de organizar las transiciones locales-globales. Dicho de otra manera, así como la percepción es imaginaria, la imaginación se basa en la percepción. Hay experimentos que demuestran que cualquier cosa que se le proporcione a un organismo (enactivamente encarnado) como excusa para una interacción sensorimotriz sirve para que el organismo inmediatamente construya un mundo que está formado, totalmente conformado. Esto representa un gigantesco paso conceptual, desde la idea que existen propiedades del mundo que hay que aprehender para formarse una imagen adecuada de la realidad, hasta la noción de que prácticamente cualquier cosa nos sirve de excusa para inventar la realidad. Nuestro mundo es imaginación y fantasía, y es por esto que es de fundamental importancia que los niños desarrollen el suyo.

El segundo corolario de este principio es que, ya que la mente se basa en la emergencia de lo local a lo global, no hay nada en la mente que podamos separar en elementos discretos y diferentes. En otras palabras, fenomenológicamente, nuestras mentes no presentan una división clara entre la memoria por un lado, y el afecto o la visión por otro. Como consecuencia, uno de los descubrimientos más impactantes de los últimos años es que el afecto o la emoción está en el origen de lo que hacemos todos los días en nuestro manejo e interacción con el mundo; que la razón o el razonamiento vendría a ser como la cereza de la torta. La razón es lo que surge en el último estadio de la emergencia minuto a minuto de la mente. Fundamentalmente, la mente es algo que emerge de la tonalidad afectiva,

que está anclada en el cuerpo. Todo el proceso ocurre en más o menos una fracción de segundo, una y otra vez. En el proceso del surgimiento momentáneo de un estado mental, los estadios tempranos están enraizados en las superfices sensorimotrices próximas a la espina dorsal en el cerebro medio, luego se dirigen hacia arriba a lo que denominamos el sistema límbico, en la corteza superior. De este modo, esta tonalidad emocional se va transformando en categorías y elementos distintos y cadenas de razonamiento, que son las clásicas descripciones de la mente en términos de unidades. Pero la razón y las categorías son literalmente los picos de la montaña que están asentados en el afecto, particularmente en el afecto y la e-moción. En efecto, la e-moción es ya intrínsicamente cognitiva. Una vez que hayan cambiado de perspectiva y dejen de considerar que la razón es el principio central de la mente, entonces podrán ver la emergencia del momento de la mente a medida que ocurre. Se origina en este caldo, el organismo total en situación, y entonces da origen a este movimiento ascendente, que se va extendiendo gradualmente como los picos de una montaña.

Es por esto que la experiencia fenomenológica es tan difícil de articular en un pie de página, ya que una parte tan considerable de su base es pre-reflexiva, afectiva, no-conceptual, pre-noética. Es difícil ponerla en palabras, justamente porque precede a la palabra. Cuando



digo que precede a la palabra no quiero decir que está más allá de las palabras. Al contrario, es porque está tan enraizada que todavía no se ha transformado en aquellos elementos de la razón que tendemos a considerar como la expresión más elevada de la mente.

Por lo tanto, para resumir ambos corolarios, en primer lugar, afirmo: la vida es como un sueño; en segundo lugar, para citar a Pascal: "El corazón tiene sus razones que la razón desconoce".

3. El punto-clave de la intersubjetividad

Con el siguiente punto clave nos adentramos en una región que todavía nos resulta bastante desconocida. El lema es: *Esta mente es* esa *mente*. Las ciencias cognitivas y del cerebro siempre han considerado que la mente está alojada en el cerebro, y, por lo tanto, que la mente del otro es impenetrable y opaca. Se considera que cualquier violación de esta separación espacial nos remite a algún tipo de energía psíquica que es necesario evitar. Pues bien, las últimas investigaciones en ciencia cognitiva nos muestran claramente que la individualidad y la intersubjetividad no se oponen, sino que son necesariamente *complementarias*. Nuevamente, no deja de asombrarme que algunos filósofos de la mente hayan gastados litros de tinta para tratar de probar que tenemos una conciencia y que no estamos rodeados de zombies. ¡Es francamente grotesco! El tema ha sido planteado al revés: la presencia y la realidad del Otro es algo tan íntimamente cercano que la pregunta pertinente es: ¿cómo pudimos haber llegado a la noción de que somos tan distintos y estamos tan separados?

Antes señalé que la investigación moderna en ciencia cognitiva nos ha proporcionado amplia evidencia de que todos los fenómenos cognitivos son también emocional-afectivos. Esto es, en forma muy natural, ha llegado ha considerar que el fundamento de la génesis de la mente es un fenómeno afectivo-empático. Esto ha quedado particularmente claro en estudios con primates superiores y niños pequeños.

Así, en un estudio reciente, Provinelli y Preuss[4] concluyen que el resultado clave de muchas décadas de investigación con primates superiores no es la controversia lenguaje/no lenguaje. Más bien, es que los primates superiores tienen una capacidad sumamente desarrollada para interpretar la *mente del otro*. Se trata de un tipo especial de inteligencia relacionada con la habilidad para reconocer estados mentales como los deseos, intenciones y creencias a partir de la presencia corporal del otro: rostro, postura y sonido. Estudios conductuales realizados con niños y chimpancés revelan impactantes similitudes en el desarrollo de ambos en relación con este punto, si bien es probable que los humanos hayan desarrollado refinamientos posteriores. En su trabajo pionero con bebés, D. Stern ya había observado que incluso en eventos perceptuales, las fronteras entre el *yo* y el otro no están claramente delimitadas, y que el ser un "yo" y constituir un "tú" son efectos concomitantes. La asombrosa capacidad del bebé para una respuesta empática surge pocas horas después del nacimiento. Un último ejemplo se relaciona con lo que todo el mundo sabe, que los niños necesitan amor y cuidado cuando son pequeños. Hace algún tiempo se publicó un estudio[5] sobre los efectos del amor y el cuidado en la estructura de los elementos locales de los bebés. Asombrosamente, se pudo demostrar que el contacto físico y los cuidados amorosos eran un determinante directo no sólo a nivel de las propiedades cerebrales (es decir, transformaciones a nivel de sinapsis y de neurotransmisores), sino también a nivel de la expresión genética. En otras palabras, estos niños son modificados en su misma constitución corporal por acciones que ocurren a nivel emocional entre seres humanos. También podríamos citar otros ejemplo, como aquellos estudios sobre la temprana infancia y el interés de los bebés por los rostros humanos. O la neurofisiología reciente que vincula la percepción de la propia imagen

---

[4] Provinelli, D.J. y T. M. Preuss (1995), Theory of mind: evolutionary history of a cognitive specialization, *Trends Neuroscience* 18:418-424.

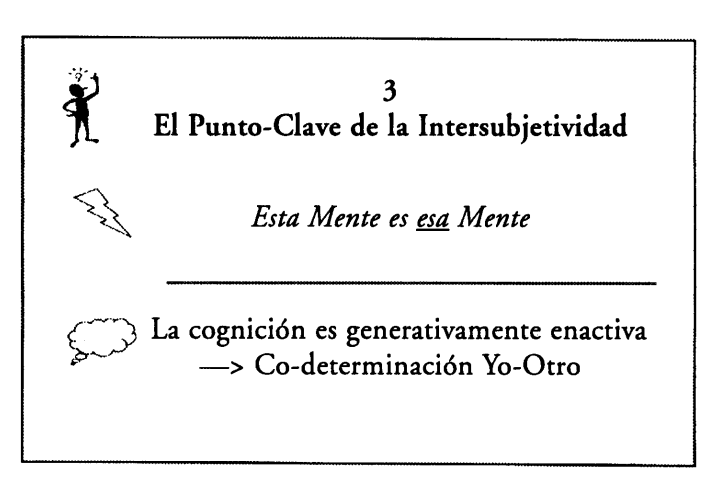
[5] Meany,M.J. et al. (1996). Early environmental regulation of forebrain glucocorticoid receptor gene expression: Implications for adrenocortical responses to stress. *Developmental Neuroscience*, 18:49-72.

corporal a la percepción del cuerpo del otro como en un espejo. El punto básico es siempre el mismo, es decir, que la cognición es *generativamente enactiva*, una codeterminación Yo-Otro.

Por supuesto, existe un vínculo directo entre el afecto y la empatía que aquí quisiera explorar un poco más. El afecto es una dinámica pre-reflexiva de autoconstitución del *self*, un autoafecto en sentido literal. El afecto es primordial, en el sentido de que soy afectado o conmovido antes de que surja un "Yo" que conoce. Mientras escribo en este preciso momento, estoy en una actitud de disposición, movido por la anticipación de escribir y plasmar mis ideas en palabras. Mientras escribo estas palabras, esta disposición se tiñe de una carga emocional que revela cierto grado de frustración por no encontrar la palabra adecuada. Pero esa tonalidad emocional surge sobre el trasfondo de un estado anímico exaltado de un día productivo dedicado a concluir este texto. Más explícitamente, quisiera distinguir tres niveles en el afecto. El primer nivel está constituido por las *emociones*: el estar consciente de una tonalidad emocional que es constitutiva del presente vivo. El segundo es el *afecto*, una disposición propia de una secuencia coherente de acciones encarnadas más largas. Finalmente, el *estado anímico*, el nivel de la narración descriptiva a lo largo de una gran extensión de tiempo (varios días o semanas).

Esta cualidad primorial o pre-verbal del afecto la vuelve inseparable de la presencia de los otros, y es aquí donde quisiera presentar el último punto de mi argumento. Para comprender por qué esto ocurre así, hay que centrarse en los correlatos *corporales* del afecto, los que no solamente aparecen como conductas externas, sino también como directamente sentidas, como parte de nuestro *cuerpo vivido*. Este aspecto de nuestro cuerpo vivido desempeña un papel decisivo en la forma en que aprehendo al Otro, no como una cosa, sino como otra subjetividad semejante a la mía, un *alter ego*. Es a través del cuerpo del otro que establezco un vínculo con el otro, primero como organismo semejante al mío, pero también percibido como presencia encarnada, lugar y medio de un campo experiencial. Esta doble dimensión del cuerpo

(orgánico/vivido; *Körper/Leib*) es un aspecto esencial de la *empatía*, el camino real para acceder a la vida social consciente, más allá de la simple interacción, como inter-subjetividad fundamental[6] .



4. El punto-clave de la circulación

El siguiente punto clave nos lleva directamente a una curva más cerrada, pero una curva necesaria, ya que también es fruto de la dinámica interna de la investigación. El lema es: *La conciencia es un asunto público.* El antecedente directo de este punto clave es el *boom* reciente del estudio de la conciencia humana como dominio legítimo de la ciencia. Pero el término conciencia es lo suficientemente vago como para justificar algunas aclaraciones previas.

La intuición que anima este punto clave es la siguiente: la profundidad inherente a la experiencia directa, vivida, permea las raíces naturales de la mente. Me gustaria desarrollar esta intuición en dos etapas. En primer lugar, un análisis científico de la mente apropiado (es decir, en el contexto de las ciencias cognitivas) conduce

[6] A estos temas (empatía e intersubjetividad) se les puede dar un riguroso fundamento filosófico; son el iceberg escondido de los últimos trabajos de Husserl, que emergen lentamente gracias a la publicación de sus estudios masivos sobre la constitución (síntesis pasiva) y la intersubjetividad (cf. *Husserliana* XI, XIV-XVI).

a la necesidad de un examen detallado de la propia experiencia, Segundo, la experiencia examinada y el análisis científico pueden tener una relación no-dual explícita, una *determinación mutua*, una circulación que evita tanto los extremos del neuro-reduccionismo y esa inefabilidad de la conciencia.

Uno de los rasgos que comparten muchas tradiciones espirituales es la noción de que la experiencia humana no debe ser tomada en su valor nominal, sino *examinada* de una u otra forma. En contraste, las ciencias cognitivas prácticamente sólo se han interesado por las facultades cognitivas en la vida común, *no*-examinada. Pero esto está cambiando rápidamente, y no es sorprendente, ya que las ciencias cognitivas se enfrentan al desafío único de incluir nuestra propia vida consciente en su ámbito de competencia, y, a fortiori, el propio acto de examinar nuestra vida individual. En otros lugares he insistido sobre la importancia de una neurociencia experimental de este tipo, o *neurofenomenología*[7]. El problema crucial de este programa de investigación es que si no queremos reducir la experiencia (examinada y no examinada) a una perspectiva neural, habrá que desarrollar una *metodología* apropiada para su examen. Es aquí donde típicamente se produce una interfase con las tradiciones espirituales. Junto al probema de la metodología surge el problema complementario de la *naturaleza* de la relación entre estos datos científicos externos, provenientes de la ciencia cognitiva, y la primera persona; esto es, datos fenomenológicos anclados en la experiencia vivida. ¿Cómo se constriñen mutuamente y se *co-determinan* estos dos ámbitos de la observación y la descripción? Es justo decir que esta línea de trabajo pone al descubierto un número de enormes y perturbadores desafíos, razón por la cual este punto clave es mucho menos consensual que los anteriores.

En relación con la metodología, ya sea fenomenológica o basada en las tradiciones contemplativas, uno de los mayores desafíos consiste en poner al descubierto cada uno de estos aspectos de la

---

[7] Varela, F. (1996), Neurophenomenology: a methodological remedy for the hard problem, *J. Consc. Studies* 3:330-350.

forma de acceder a fenómenos en primera persona, en carne y hueso, para establecer una *pragmática* fenomenológica más allá de un uso vago. El fracaso por transformar la reducción fenomenológica en un método concreto es, a mi juicio, el tema menos desarrollado de la Fenomenología. Como resultado, una cantidad apreciable de literatura fenomenológica ha derivado en un análisis textual y en repeticiones de descripciones de Husserl u otros, sin el compromiso disciplinado de re-hacer nuevamente estas descripciones. En occidente, más allá de la fenomenología, están las muy criticadas escuelas introspeccionistas de los inicios de la psicología del siglo XX que en este momento están en proceso de revisión[8].

Las prácticas de esta experiencia de examen revelan *una región claramente determinada de reciprocidad ontológica.* Esta región puede ser caracterizada por la forma en que entidades específicas surgen como desde adentro (o sea, lo que está *dado*), y por su clara constitución inmanente (es decir, lo que está dado está dado en una esfera que es *mía*). Esta región específica está habitada o revela entidades y eventos inseparables de la presencia del cuerpo vivido en toda su complejidad. No necesito decir que la línea de trabajo que acabo de esbozar no puede ser desarrollada sobre una base abstracta y general, sino que tiene que basarse en estudios de caso, realizados paso a paso. Como ejemplo, podemos mencionar el estudio del problema central de la neuro-fenomenología de la experiencia del tiempo, el siempre presente ahora encarnado[9], o el origen de la imagen del *yo*.

Así, el estudio de la conciencia en el sentido que aquí examinamos abre ontologías regionales de objetos materiales o ideales que están inter-penetrados, transparentes tanto hacia su base material (de otro modo permanecerían inexorablemente separados) y al ámbito de la experiencia (de otro modo serían

---

[8] P. Vermersch, "Introspection as Practice", en F. Varela y J. Shear, *The View from Within: First-person methods in the exploration of consciousness, J. Consc. Studies* 6(2-3), 1999.

[9] F. Varela, "The Specious Present; the neurophenomenology of time consciousness", en J. Petito, F. Varela, B. Pachoud, y J.M. Roy (Eds.), *Naturalized Phenomenology*, Stanford University Press, Stanford.

idealidades flotantes desencarnadas). Claramente, la reciprocidad *Leib/Körper* no estará disponible en toda su enorme importancia mientras conservemos una actitud de experiencia no examinada, esto es, una actitud natural. Si queremos avanzar más lejos, habrá que cultivar la actitud reductiva.

Un punto importante es que la neurofenomenología nos ha internado por los caminos de la renovación filosófica y metodológica. Si pretendemos que esta línea de investigación proporcione una respuesta para lo que de otra manera sería una brecha explicativa insalvable entre la mente cognitiva y la mente fenomenológica, no podemos ignorar la base constitutiva de la reciprocidad mutua que hace que lo mental y lo experiencial, lo corporal y lo neural, se mantegan unidos. Por lo tanto, es evidente que es sólo a partir de esta base renovada que la neurofenomenología puede llegar a ser algo más que una repetición del pasado, bajo la forma de búsquedas de correspondencia que atraviesan la línea "de misterio".

Esta reciprocidad mutua sin residuo constituye la naturaleza misma de esta región que es propia del *Körperleib*. Al final, para que esta región ontológica se manifieste en toda su luminosidad, habrá que enlazar tres elementos para entretejer una urdimbre de continuidad entre lo material y lo experiencial, lo natural y lo trascendental[10]. En otras palabras, hemos identificado tres polos en la circulación mutua que hemos estado examinando:

(I) el nivel *formal* ya que la descripción de los contenidos mentales participa de un modo de idealidad y, por lo tanto, están efectivamente en terreno común;

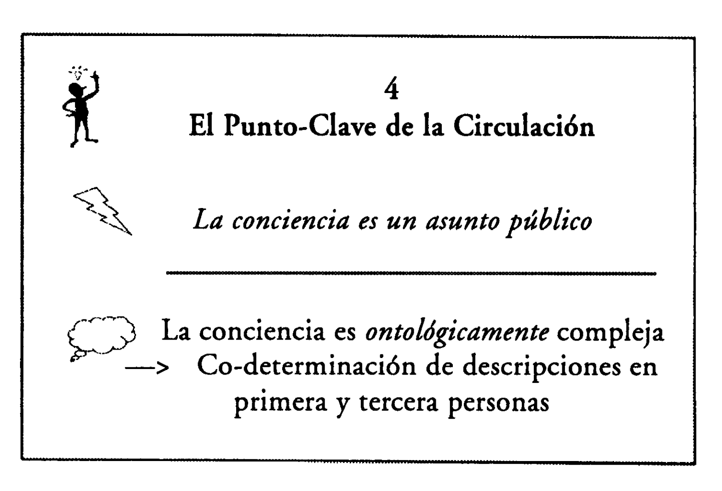
(II) el proceso *natural* (neural, corporal) considerado en su nivel apropiado, tendiendo un puente entre la emergencia global y los mecanismos locales, que le dan una relevancia directa tanto al contenido psicológico y a un examen neurocientífico detallado;

---

[10] F. Varela (1997), The Naturalization of Phenomenology as the Transcendence of Nature: Searching for generative mutual constraints, *Alter: Revue de Phénoménologie Nº 5* pp. 355-385 (París).

(III) el nivel *pragmático* de examen que conduce a la transición *Leib/Körper* ya que sólo este nivel nos permite acceder a una posición no-dual, que no excluye a la experiencia ni al cuerpo, y proporciona la base relevante o los datos para I-II.

Entre estos tres polos no hay simplemente una relación estática o estructural: están en una relación mutuamente generativa en el sentido de que cada uno requiere del otro para cobrar sentido. Ninguno de ellos se basta a sí mismo en forma aislada. Creo que esta triple urdimbre puede entregarnos los fundamentos para una renovación importante en filosofía y en ciencia, permitiéndonos acceder a un pensamiento no dual, no declarativo ni pre-definido por decreto, sino que se halla en el umbral de nuestra propia puerta.



NEUROFENOMENOLOGÍA

Un remedio metodológico para el difícil problema

La unión de estos dos términos —neurología y fenomenología— es, en gran medida, una provocación premeditada. Este trabajo es un intento decidido (y un tanto radical) de encontrar una salida a la estéril oposición entre los fundamentos biológicos de la conciencia y el hecho básico de tener una experiencia irreductible, lo que se ha dado en llamar el "difícil" problema de la conciencia, siguiendo al filósofo David Chalmers. Propongo aquí la disolución del problema a partir de un ángulo metodológico donde una pragmática en primera persona (es decir, lo que se experimenta directamente, y no como hecho de observación externa) es lo central, y a la que la tradición fenomenológica continental da sus bases.

Este artículo responde a los problemas planteados por D.J. Chalmers (1995), ofreciendo una dirección de investigación que es bastante radical en el sentido en que algunos principios metodológicos básicos se relacionan con los estudios científicos de la conciencia. Neuro-fenomenología es el nombre que utilizo aquí para designar una búsqueda para unir la ciencia cognitiva moderna con un *enfoque disciplinado* de la experiencia humana, poniéndome así en el linaje de la tradición continental de la fenomenología[1].

(Varela, F. (1996), Neurophenomenology: A Methodological remedy to the hard problem, *J.Consc.Studies* 3:330-350).

1 La utilización de "neuro" debe ser tomada aquí como un *nom de guerre*. El término es elegido en evidente contraste con el uso común del término "neurofilosofía", que identifica a la filosofía con la filosofía angloamericana de la mente. Además, "neuro" remite aquí al despliegue total de correlatos científicos que son pertinentes en la ciencia cognitiva. Sin embargo, sería excesivamente engorroso hablar de una fenomenología-neuro-psico-evolucionista.

Afirmo que el denominado difícil problema que anima estas Ediciones Especiales del *Journal of Consciousness Studies* sólo puede ser dirigido de manera productiva reuniendo a una comunidad de investigación armada con nuevas herramientas pragmáticas que le permitan desarrollar una ciencia de la conciencia. Diré que ninguno de los correlatos empíricos poco sistemáticos, ni principios puramente teóricos, nos ayudarán realmente en esta etapa. Necesitamos cambiar de rumbo hacia un estudio sistemático del único vínculo entre la mente y la conciencia que parece obvio y natural, a la vez: la *estructura de la experiencia humana misma.*

A continuación, comenzaré mi propuesta examinando brevemente el debate actual acerca de la conciencia a la luz del "difícil problema" de Chalmers. Después, esbozaré la estrategia (neuro)-fenomenológica. Concluiré analizando algunas de las principales dificultades y consecuencias de esta estrategia.

1. Una cartografía de enfoques

El enigma de la experiencia

Chalmers empieza el debate del "difícil problema" centrándose en *el* problema que parece fundamental: la *experiencia* relacionada con los sucesos cognitivos o mentales.

A veces, términos como "conciencia fenomenal" y "*qualia*" también son utilizados aquí, pero creo que es más natural hablar de "experiencia consciente" o simplemente de "experiencia" (Chalmers, 1995, p.201).

Después de describir estudios de caso de algunas explicaciones funcionalistas populares, Chalmers pasa a calificar el desafío restante como un "ingrediente extra" necesario. Ya la elección del término es reveladora, puesto que Chalmers parece asumir desde el principio que el único camino es encontrar principios teóricos que salven la distancia entre la cognición y la experiencia. Tal como expondré más adelante, parece que otra alternativa fundamental es cambiar todo el marco dentro del cual el tema es tratado. De

cualquier modo, "la moraleja de todo esto es que no *se puede explicar la experiencia consciente a bajo precio*" (p.208; en itálicas en el original). Estoy completamente de acuerdo, pero me apresuro a decir que el precio que necesitamos pagar es mucho mayor que aquel que la mayoría de las personas está dispuesta a conceder. Una vez más, la principal dificultad es que la experiencia no constituye "un planteamiento explicativo, sino un *explanandum* por derecho propio y, por lo tanto, no es una candidata para la eliminación [reduccionista]" (p.209). Lo que se necesita, concluye, es un tipo de explicación no-reduccionista. Una vez más, estoy de acuerdo con Chalmers, pero una de mis tareas será exponer en detalle cuán diferentes son nuestras opiniones a partir de este punto.

Permítanme comenzar mi nuevo enfoque sobre el problema de la experiencia en el actual auge en el estudio científico de la conciencia. Como todos sabemos, la cantidad de libros, artículos y reuniones acerca del tema ha aumentado de manera exponencial en los últimos años. ¿Por qué este arrebato actual después de todos los años de silencio, en los cuales la conciencia era un tema "mal visto", incluso en la ciencia cognitiva?

Por cierto, después del punto máximo de dominio del *behaviorismo* debió existir una fase conservadora antes de que la ciencia cognitiva sintiera que estaba pisando suelo firme. Quizá fue más importante el estilo de la filosofía de la mente, dominante en los EE.UU (con muchos seguidores en Europa), que desconfía intrínsecamente de la experiencia subjetiva. En este marco, se han realizado avances importantes en la ciencia cognitiva casi exclusivamente dentro de una perspectiva cognitivista-computacionalista o conexionista. El conexionismo, en particular, hizo posible una idea revolucionaria de transiciones y puentes entre niveles de explicación, mejor entendida como una filosofía de la emergencia: cómo las reglas locales pueden originar objetivos o propiedades globales en una causalidad recíproca. Esto dio un nuevo significado a la tradicional interfaz mente/cuerpo, que en forma de procesos cognitivos, como esquemas computacionalistas o conexionistas, hizo una selección de los fenómenos cognitivos específicos (excelentes ejemplos son la visión, el movimiento y la memoria asociativa) que, en principio, tienen solución (por no decir que están

solucionados). Al mismo tiempo, estos avances crearon el mismo trasfondo para el "difícil problema", puesto que hicieron aparecer a la conciencia como desprovista de cualquier significación causal. Esto está bien ilustrado en el libro innovador de Ray Jackendoff, en que la "mente fenomenológica" (a saber, conciencia qua experiencia) es vista como la proyección de una "mente computacional" (a saber, mecanismos cognitivos) en que toda causalidad tiene lugar. De este modo, la única conclusión a la que pudo llegar es que la conciencia "no es buena para nada" (1987, p.26).

Además, en avances paralelos, las nuevas técnicas para análisis a gran escala de la actividad cerebral y la neurosicología nos han permitido, por primera vez, hacer preguntas experimentales directas acerca de los correlatos de cognición compleja en acción, tales como las imágenes mentales y las emociones (ver por ejemplo, Posner y Raichle, 1992; Mazoyer, Roland y Fox, 1995). Los experimentos que intervienen en estas mediciones no-invasivas en línea son particularmente interesantes puesto que han llevado a los investigadores a enfrentar preguntas como: ¿Puede creerse en el informe de un sujeto? ¿Cuáles son las expresiones verbales del informe? Estas son preguntas básicamente experienciales, que ya implican una revisión importante de la forma en que los informes de la experiencia humana deben ser abordados en la investigación empírica.

Un día la historia intelectual de los característicos giros y vueltas de este problema será reexaminada rigurosamente. Pero éste tiene un aura de *déjà-vu*, que nos hace recordar las muchas oscilaciones del péndulo, entre el rechazo y la total fascinación con los debates científicos acerca de la experiencia consciente. Esto no podría ser de otra manera, ya que toda ciencia de la cognición y de la mente debe, tarde o temprano, aprender la condición básica de que, aparte de nuestra experiencia de ello, no sabemos qué podría ser lo mental o lo cognitivo. Como John Searle ha observado acertadamente en su propia contribución al auge, si existe una fase de investigación que favorezca a las teorías de la mente estrictamente materialistas:

[El filósofo] encuentra dificultades. Siempre parece que está dejando algo fuera... [y] bajo las objeciones técnicas subyace una

objeción mucho más profunda... [esto] puede ser dicho de manera bastante simple: La teoría en cuestión ha dejado fuera a la mente; ha excluido alguna característica esencial de la mente, como la "conciencia" o la "*qualia*" o el contenido semántico... [Por consiguiente], si tuviéramos que pensar en la filosofía de la mente como un solo individuo, diríamos que esta persona es un neurótico compulsivo y su neurosis toma la forma de repetir el mismo patrón de conducta una y otra vez (Searle, 1992, p. 30-1).

Estoy tan de acuerdo con el diagnóstico así como en desacuerdo con la cura propuesta por Searle (más tarde diré algo más de esto). Claramente, necesitamos algunas medidas radicales para compensar esta conducta compulsiva. Esto es precisamente lo que quiero hacer aquí, con una propuesta que probablemente parecerá radical a algunos: pero sólo con ella se romperá el círculo vicioso y se evitarán los intentos de fijarla con otro modelo teórico abstracto.

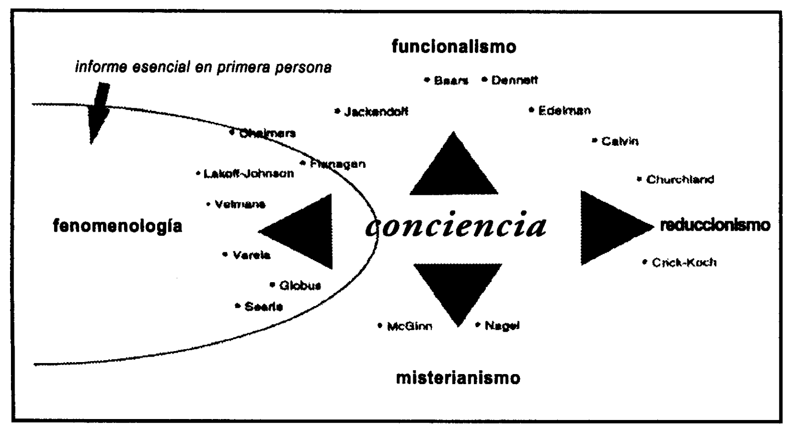


Figura 1

Un esquema de cuatro vías

Con el fin de comprender mi postura, ahora el lector debería mirar el esquema de la Figura 1 que esboza cuatro ejes que parecen reproducir las orientaciones esenciales del auge actual del debate acerca de la conciencia. No pretende ser un gráfico de los diversos puntos de vista

que abarque todo, sino una oportunidad para ponerme en contexto con los autores modernos que han publicado argumentos extensos (generalmente en formato de libro) en los últimos años[2].

A la derecha, he colocado la misma tendencia que se hace oír mejor representada por P.S. Churchland, pero incluyendo a F. Crick y C. Koch, y cerca de la filosofía espontánea de muchos otros de mis colegas de neurociencia, apropiadamente clasificada como neuro-reduccionismo o eliminativismo. Como es bien sabido, esta perspectiva busca resolver el difícil problema a través de la eliminación del polo de experiencia en favor de alguna forma de explicación neurobiológica, que haría el trabajo de generarlo (Churchland y Sejnowski, 1992). O como dijo Crick con su característica franqueza: "Uno no es nada más que un paquete de neuronas" (1994, p.2).

En la parte superior del centro, he reunido una variedad de posiciones que pueden ser clasificadas como funcionalistas, e identificadas por Chalmers como la ecología de ideas más popular activa hoy (1995, p. 204-9). En los últimos 20 años, se ha preferido drásticamente el funcionalismo en la ciencia cognitiva, seguido por la estrategia de reemplazar el vínculo entre la cognición y la conciencia (el más inmediato en la tradición filosófica occidental) por el vínculo entre la cognición y sus correspondientes estados funcionales o intencionales. En el mejor de los casos, el problema de la conciencia es asimilado con el de "*qualia*" para algunas características particulares de los estados mentales. Así, la noción de experiencia llega a ser convincentemente asimilada con la de conducta cognitiva, actitud proposicional o papel funcional.

Estas perspectivas incluyen una serie de propuestas muy desarrolladas que comprenden el "Mecanismo Proyectivo" de R. Jackendoff (1987), el "Espacio de Trabajo Global" de B. Baars (1988),

2 Obsérvese que este es un gráfico de enfoques naturalistas, es decir, de posiciones que, cada una a su manera, proporcionan un vínculo factible a la investigación actual acerca de la ciencia cognitiva. Esto excluye al menos dos corrientes de debate popular: por una parte, visiones que adoptan una postura dualista tradicional (à *la* J.C. Eccles); por otra parte, exigencias de nuevos fundamentos de los defensores de los mecanismos cuánticos. Creo que ambas visiones son extremas e innecesarias y se concentran en aquellos que se basan, de manera explícita, en la ciencia cognitiva y la neurociencia actual.

las "Versiones Múltiples" de D. Dennett (1991), las "Máquinas Darvinianas" de W. Calvin (1990) o el "Darwinismo Neural" de G. Edelman (1989). El movimiento básico de estas proposiciones es bastante similar. En primer lugar, comienzan desde los puntos modulares de las capacidades cognitivas (a saber, los problemas "fáciles"). En segundo lugar, crean un marco teórico para juntarlos de manera que su unidad equivale a una explicación de la experiencia. La estrategia para construir esta unidad emergente y la experiencia misma varían, pero es característico que se deje imprecisa, puesto que todo el enfoque depende casi completamente de una tercera persona o de un enfoque externo para obtener información y validar la teoría. Esta parece ser la posición más popular en las publicaciones actuales en auge, y representa a un segmento importante de investigadores de la ciencia cognitiva. Esta popularidad se basa en la aceptación de la realidad de la experiencia y de la vida mental mientras se mantienen los métodos y las ideas dentro del conocido marco de la ciencia empírica.

En la parte de abajo del centro tenemos el reflejo del funcionalismo. Los misterianistas como T. Nagel (1986) y C. McGinn (1991) intentan concluir mediante argumentos de principios que el difícil problema no tiene solución, basándose en limitaciones intrínsecas de los medios a través de los cuales adquirimos nuestro conocimiento de lo mental.

Finalmente, a la izquierda, coloqué el sector que más me interesa, y que puede ser descrito, en líneas generales, como el que da un papel central y *explícito* a los informes en primera persona y a la naturaleza irreductible de la experiencia, mientras que al mismo tiempo rehúsa tanto una concesión dualística como una renuncia pesimista al problema, como en el caso de los misterianistas. Esto sigue la línea de la identificación de Chalmers de donde se encuentra el difícil problema. Tal como en las otras orientaciones de mi esquema, el grupo reunido aquí es heterogéneo, con extrañas combinaciones, como el enfoque de la semántica cognitiva de G. Lakoff y M. Johnson (1987), las ideas sobre la irreductibilidad ontológica de J. Searle (1992), el cerebro "postmoderno" de G. Globus (1995) y al final, "el equilibrio reflexivo" de O. Flanagan (1992) y

la propia propuesta de Chalmers tal como la desarrolló en detalle en su reciente libro.

Lo que es interesante acerca de este variado grupo, dentro del que me incluyo, es que, aun cuando compartimos un interés por la experiencia directa como un hecho básico para incorporar en el futuro de la disciplina, son patentes las diferencias en la forma en que se toma en cuenta la experiencia. El enfoque fenomenológico se basa en un paso característico hacia el estudio de la experiencia, que es el núcleo de mi propuesta. Espero que esto aclare lo suficiente el contexto de mis ideas dentro de la escena actual. Ahora podemos pasar al centro de la cuestión, la naturaleza de la circulación entre un informe en primera persona y un informe externo de la experiencia humana, lo cual describe la posición fenomenológica en un diálogo fértil con la ciencia cognitiva.

2. Un enfoque fenomenológico

La irreductibilidad: lo fundamental

El enfoque fenomenológico comienza a partir de la naturaleza *irreductible* de la experiencia consciente. La experiencia vivida es desde donde empezamos y con lo que todos volvemos a relacionarnos, como un hilo conductor. La mayoría de los autores modernos no se sienten inclinados a concentrarse en la distinción entre la vida mental, en un sentido general, y la experiencia, o manifiestan alguna desconfianza acerca de su estatus.

Desde un punto de vista fenomenológico, la experiencia consciente discrepa bastante de la experiencia de contenido mental, tal como figura en la filosofía angloamericana de la mente. La tensión entre estas dos orientaciones aparece en forma algo espectacular en el libro de Dennett (1991), donde concluye con poco esfuerzo (15 líneas en un libro de 550 páginas) que la fenomenología ha fracasado. Señala:

> *"Como otros intentos para sacar una interpretación y revelar los hechos básicos de la conciencia para una observación rigurosa, tal como los movimientos impresionistas en las artes [*sic*] y los sicólogos*

introspeccionistas como Wundt, Titchener y otros, la fenomenología ha fracasado en su intento por encontrar un único método ordenado con el cual todos pudieran estar de acuerdo" (p.44).

Este pasaje es revelador: Dennett mezcla peras con manzanas al colocar en el mismo saco el impresionismo y el introspeccionismo; confunde introspeccionismo con fenomenología, que definitivamente no se asemejan (*vide infra*); y finalmente saca su conclusión de la ausencia de un acuerdo idílico que validaría el todo. Bueno, nosotros no pedimos que "todo el mundo estuviera de acuerdo", digamos, con el darwinismo, para hacer de él un programa de investigación sorprendentemente útil. Y ciertamente *algunas* personas están de acuerdo con la sólida posibilidad de un examen disciplinado de la experiencia humana. En un libro que es tan *savant* y perspicaz en muchos otros aspectos, esta demostración de ignorancia en lo que respecta a la fenomenología es un síntoma que dice mucho acerca de cuál es el problema en este ámbito.

El punto principal que debe ser llevado a primer plano es claramente planteado por Searle:

> *"...la mayor parte del fracaso de gran parte del trabajo realizado en la filosofía de la mente... en los últimos cincuenta años... proviene de un fracaso constante en reconocer y aceptar el hecho de que la ontología de lo mental es una ontología irreductible de primera persona... En resumen, para nosotros no existe forma de imaginar la subjetividad como parte de nuestra visión de mundo, porque, por decirlo así, la subjetividad en cuestión es la imagen" (Searle, 1992, p. 95, 98).*

Sin embargo, en la defensa de la irreductibilidad de la conciencia de Searle existe una incapacidad para llegar a cualquier conclusión acerca de cómo resolver el problema epistemológico con respecto al estudio de la conciencia. Searle quiere que aceptemos que "la irreductibilidad de la conciencia es sólo una consecuencia de la pragmática de nuestras prácticas definicionales" (p. 122), y, por lo tanto, aunque la irreductibilidad de la conciencia es un "argumento sencillo", "no tiene consecuencias profundas" (p. 118). De hecho,

El hecho mismo de la subjetividad, que estábamos tratando de observar, hace que dicha observación sea imposible. ¿Por qué? Porque donde está involucrada la subjetividad consciente, no existe distinción entre el observador y la cosa observada... Cualquier introspección que haga de mi propio estado consciente es en sí mismo ese estado consciente (p.97).

Lo mental no tiene una forma clara de investigarse a sí mismo, y se nos deja con una clara conclusión lógica, pero en un limbo pragmático y metodológico.

Esto no se diferencia del limbo en las perspectivas de Jackendoff, quien, a su manera, afirma la irreductibilidad de la conciencia, pero es reveladoramente silencioso cuando llega al método. Él afirma que las percepciones en la experiencia actúan como restricciones para una teoría computacional de la mente, pero continúa sin ninguna recomendación metodológica a excepción de "la esperanza de que las discrepancias sobre la fenomenología puedan ser establecidas en una atmósfera de confianza mutua" (Jackendoff, 1987, p. 275). ¡Confianza mutua! Lo que se necesita es un mét*odo* riguroso y aquí reside tanto la dificultad como el potencial revolucionario del tema.

Método: tomar la delantera

Necesitamos examinar, más allá del fantasma de la subjetividad, las posibilidades concretas de un examen disciplinado de la experiencia que se encuentran en el núcleo mismo de la inspiración fenomenológica. Para repetir: este es el redescubrimiento de la primacía de la experiencia humana y su cualidad vivida directa, que constituye el proyecto de base de la fenomenología. Este es el sentido con el cual Edmund Husserl inauguró su pensamiento en Occidente y estableció una larga tradición que hoy está bien y viva no sólo en Europa, sino en todo el mundo. De hecho, entre 1910 y 1912, mientras Husserl estaba en el momento máximo de su creativa formulación de la fenomenología, en los Estados Unidos, William James estaba siguiendo líneas muy paralelas en su enfoque

pragmático a la vida cognitiva. Y para completar la "sincronía" planetaria de este giro, apareció en Japón una renovación filosófica muy innovadora, la denominada Escuela de Kyoto, iniciada por Nishida Kitaro y continuada por Nishitani Keiji y otros. Husserl y James se conocieron y se leyeron el uno al otro y los miembros de la escuela de Kyoto leyeron mucho acerca de la fenomenología occidental y pasaron largos períodos de formación en Alemania. Así, creo que deberíamos considerar estos *anni mirabiles* para la fenomenología, similares a los años 1848-52 para el nacimiento de la biología evolucionista moderna.

Es justo decir que la fenomenología es, más que cualquier otra cosa, un *estilo de pensamiento*, que Husserl comenzó en Occidente, pero que no es agotado por su estilo y opciones personales (Lyotard, 1954). No quiero entrar en una explicación de la diversidad y complejidad de la fenomenología occidental (ver, por ejemplo, Spiegelberg, 1962). Las contribuciones de individuos como Eugen Fink, Edith Stein, Roman Ingarten y Maurice Merleau-Ponty, para citar sólo a unos pocos, han atestiguado un continuo desarrollo de la fenomenología. Más recientemente, se han estudiado varios vínculos con la ciencia cognitiva moderna (ver, por ejemplo, Dreyfus, 1982; Varela *et al.* , 1991; Klein y Wescott, 1994; Petitot, 1995; Petitot *et al.*, 1996; Thompson y Varela, 1996). Menciono explícitamente esto porque he observado que la mayoría de las personas que no están familiarizadas con el movimiento fenomenológico asumen automáticamente que la fenomenología es una especie de escolasticismo husserliano, algo que mejor es dejado a los polvorientos filósofos continentales que pueden leer alemán.

Como mucho, los científicos cognitivos podrían haber leído la colección editada por Dreyfus (1982), que presenta a Husserl como un tipo de proto-computacionalista, y suponen que este pedazo de la historia es todo lo que hay que saber acerca de la fenomenología. Ésta se ha transformado en una interpretación muy citada, pero los críticos han aclarado que la lectura cognitiva de Husserl por parte de Dreyfus está muy viciada. Esta no es la ocasión para extenderse en este tema, pero es esencial señalar una advertencia

aquí, no sea que el lector con una formación científica piense que esto ha sido establecido de una vez y para siempre[3].

Mi postura no puede ser atribuida a ninguna escuela particular o sub-linaje, sino que representa mi propia síntesis de la fenomenología a la luz de la ciencia cognitiva y otras tradiciones que se centran en la experiencia humana. También se puede describir la fenomenología como u*n tipo especial de reflexión* o actitud acerca de nuestra capacidad de estar conscientes. Toda reflexión revela una variedad de contenido mental (actos mentales) y su orientación correlacionada o contenidos intencionados. La actitud natural o ingenua supone una cantidad de afirmaciones recibidas acerca de la naturaleza del experimentador y sus objetivos buscados. El punto Arquímedeano de la fenomenología consiste en la suspensión de estas afirmaciones comunes y la catalización de un examen nuevo. De ahí la famosa máxima de Husserl: "¡De vuelta a las cosas mismas!",[4] que para él significaba —lo contrario de una objetificación de la tercera persona— un retorno al mundo tal como es experimentado en su inmediatez sentida. La esperanza de Husserl y la inspiración básica de la investigación fenomenológica era que una verdadera ciencia de la experiencia fuera gradualmente establecida, que no sólo podría estar en iguales condiciones que las ciencias naturales, sino que, de hecho, les otorgara la base que necesitan, porque todo conocimiento emerge necesariamente de nuestra experiencia vivida.

Por una parte, la experiencia está invadida de comprensiones previas espontáneas, de manera que podría parecer que toda "teoría" acerca de ella es bastante superflua. Pero por otro lado, esta comprensión previa debe ser examinada porque no está claro con qué clase de conocimiento se identifica. La experiencia exige un examen específico con el fin de liberarla de su estatus de creencia común. Como lo expresa Merleau-Ponty:

---

3 Para una crítica de Dreyfus, acerca de Husserl, ver Langsdorf (1985) y también las objeciones de R. McIntyre (1986). Para una explicación reciente acerca de esta controversia mediante un contraste entre Fodor y Husserl ver J.M. Roy (1995).

4 "Zurück zu den Sachen selbst!", *Logische Untersuchungen*, Vol. II, Parte 1, p.6.

*"Volver a las cosas mismas es volver a ese mundo que precede al conocimiento, del cual el conocimiento siempre h*abla *y en relación con el cual toda esquematización científica es un lenguaje de símbolos derivado y abstracto, tal como la geografía lo sería con respecto a un bosque, una pradera, un río en el campo que conocemos de antemano" (M. Merleau-Ponty, 1945, p.ix).*

Insisto en traer a primer plano este principio básico del enfoque fenomenológico, puesto que a menudo se traduce rápidamente como una búsqueda empírica de correlatos mentales. Necesitamos volver repetidamente a este tema, porque sólo apreciando su profundidad los puentes fenomenológicos pueden afirmar la mantención de un vínculo importante con la experiencia vivida y ser un remedio para el difícil problema.

La fenomenología funda su movimiento en dirección a una nueva mirada hacia la experiencia en un gesto de reflexión específico o *reducción fenomenológica* (RF)[5]. Ahora necesito desplegar en profundidad esta actitud o gesto, que es la forma habitual que tenemos para relacionarnos con los cambios de nuestro mundo vivido. Esto no requiere que imaginemos un mundo diferente, sino más bien considerar al presente como *otro.* Como dijimos antes, este gesto transforma una experiencia simple o sin examinar en una experiencia reflexiva o de segundo orden. La fenomenología insiste correctamente en este cambio desde lo natural a la actitud fenomenológica, puesto que sólo entonces el mundo y mi experiencia aparecen abiertos y necesitan ser explorados. El significado y la pragmática de la RF han tomado muchas variantes de este tronco común. No es mi intención resumirlas aquí[6].

5 El lector debería resistir la tentación de asimilar este uso de la palabra "reducción" con la "reducción teórica" tal como, por ejemplo, aparece en el marco neurorreduccionista y bien articulada en los escritos de P.S.Churchland. Los dos significados se oponen completamente el uno al otro; por lo tanto es conveniente añadir un calificador.

6 Para un debate reciente acerca de las variedades de reducción ver: R. Bernet (1994), p. 5-36. La primera articulación de Husserl puede encontrarse en sus lecturas avanzadas de 1907 (Husserl, 1962).

El gesto consciente que es el fundamento de la RF puede ser analizado en cuatro momentos o aspectos entrelazados:

(1) Actitud: reducción

La actitud de la reducción es el punto de partida necesario. También puede definirse por sus similitudes con la duda: una suspensión repentina y momentánea de las creencias acerca de lo que está siendo examinado, un poner en desuso nuestro discurso habitual acerca de algo, un poner entre paréntesis la estructura preprogramada que constituye el contexto omnipresente de la vida diaria. La reducción es autoinducida (es un gesto activo) y busca ser resuelta (disipando nuestras dudas), puesto que está aquí como una fuente de experiencia. Es un error común suponer que suspender nuestro pensamiento habitual significa detener el flujo de pensamientos, lo cual no es posible. El punto es devolver la dirección del movimiento del pensamiento de su dirección habitual orientada al contenido hacia el surgimiento de los pensamientos mismos. Esto es ni más ni menos que la misma capacidad humana para la *reflexibilidad*, y el alma de la reducción. Dedicarse a la reducción es cultivar una capacidad sistémica para la reflexión en el lugar, abriendo así nuevas posibilidades dentro de nuestro flujo mental habitual. Por ejemplo, justo ahora es muy probable que el lector esté haciendo algunas observaciones internas con respecto a qué es la reducción, qué le recuerda y así sucesivamente. La movilización de una actitud de reducción comenzaría por notar esos patrones automáticos de pensamiento, dejarlos fluir y hacer volver la reflexión hacia su fuente.

(2) Intimidad: intuición

El resultado de la reducción es que aparece un ámbito de experiencia menos obstruido y más vívidamente presente como si estuviera sin la niebla característica que separa al experimentador y al mundo. Tal como lo observó William James, así, la inmediatez de la experiencia aparece rodeada de una diversidad de horizontes hacia los cuales podemos volver nuestro interés. Este aumento de intimidad con el fenómeno es crucial, porque es la base del criterio

de verdad en el análisis fenomenológico, la naturaleza de su evidencia. Si bien la intimidad o inmediatez es el comienzo de este proceso, éste continúa mediante el cultivo de *variaciones* imaginarias, considerando en el espacio virtual de la mente múltiples posibilidades del fenómeno a medida que éste aparece. Estas variaciones ideales nos son familiares a partir de las matemáticas, pero aquí están puestas al servicio de cualquier cosa que se transforme en el centro de nuestro análisis: la percepción de la forma tridimensional, la estructura del "ahora", las manifestaciones de empatía, etc. Es a través de estas múltiples variaciones que surge una nueva fase de comprensión, una experiencia "¡Ajá!" que añade nueva evidencia que lleva una fuerte convicción. Esta intimidad que se mueve con nuestra experiencia equivale bien a lo que se alude tradicionalmente como intuición, y representa, junto con la reflexión, las dos capacidades humanas principales que son movilizadas y cultivadas en la RF.

(3) Descripción: invariantes

Detener la reducción seguida por variaciones imaginarias sería condenar este método a la determinación privada. El próximo componente es tan importante como los precedentes: el aumento de la evidencia intuitiva debe ser inscrito o traducido a artículos comunicables, generalmente, a través de lenguaje u otras inscripciones simbólicas (piénsese en esquemas o fórmulas). Sin embargo, las materialidades de estas descripciones también son una parte constitutiva de la RF y dan forma a nuestra experiencia tanto como la intuición que las constituye. En otras palabras, no estamos hablando sólo de una "codificación" en un registro público, sino más bien de una "encarnación" que encarna y da forma a lo que experimentamos. Quiero referirme a estas descripciones públicas como *invariantes*, puesto que es a través de las "variaciones" que uno encuentra numerosas condiciones bajo las cuales una observación puede ser comunicable. Esto no es muy diferente de lo que los matemáticos han hecho por siglos: la novedad es aplicarlo a los contenidos de la conciencia.

(4) Formación: estabilidad

Como en toda disciplina, la clave es la formación continua y el aprendizaje constante. Una revisión superficial de la conciencia es muy diferente del cultivo de la RF. Este punto es particularmente importante aquí, puesto que la actitud de reducción es evidentemente frágil. Ningún estudio sistemático puede madurar si uno no cultiva la *habilidad* de estabilizar y profundizar la capacidad de intuición y de poner entre paréntesis estando atento, así como la habilidad para realizar descripciones esclarecedoras. Este último aspecto de la RF es quizá el mayor obstáculo para la constitución de un programa de investigación, puesto que esto implica un compromiso disciplinado por parte de una comunidad de científicos (más adelante hablaré más acerca de esto).

Reducción Fenomenológica	
aspectos del método	***Características del examen resultante***
Actitud	poner entre paréntesis, suspensión de las creencias
Intuición	intimidad, evidencia inmediata
Invariantes	inscripciones, intersubjetividad
Formación	estabilidad, pragmática

Evitar algunas trampas comunes

En las presentaciones anteriores de estas ideas he encontrado una cantidad recurrente de malentendidos y conclusiones erróneas. Permítanme anticipar algunas de esas trampas comunes y tratarlas inmediatamente.

El análisis fenomenológico no sólo es introspeccionismo

Como muchos han notado, la introspección presupone que tenemos acceso a nuestra experiencia de la misma forma en que tenemos acceso a un campo visual "interno", como lo sugiere la etimología de la palabra, mediante una inspección. Este examen interno es una aptitud cognitiva normal de doblaje reflexivo, un gesto en el que nos ocupamos regularmente. Éste supone algún "yo" referente que hace la autoobservación, una red narrativa que da forma a lo que identificamos como un sujeto.

En los días previos a la fenomenología (a saber, sin reducción) la introspección provocó una ola de interés en la psicología, comenzando con el trabajo de W. Wundt, seguido por otros como E. Tichener en Estados Unidos y la escuela de Würzburg. A pesar de un entusiasmo inicial, el programa de investigación propuesto por el introspeccionismo no se arraigó. Entre otros problemas, los informes de diversos laboratorios no pudieron alcanzar una base común de validación. Un caso clásico fue el problema de si las imágenes visuales cumplían o no un papel en la solución de problemas. El método empleado comenzó con la reflexión, pero dio una dirección explícita en cuanto a qué buscar o qué clase de distinciones hacer, más de lo que estamos acostumbrados a ver en la sicología experimental moderna. Inevitablemente, los informes estuvieron cada vez más influenciados por los respaldos teóricos de los estudios, y, de hecho, degeneraron en discusiones acerca de la autoridad. El histórico informe de Lyons (1986) es escrito como un obituario para la introspección. Pero, tal como nos lo recuerda Howe (1991), esta sería una conclusión prematura.

Para un fenomenólogo, esta forma de movilizar capacidades reflexivas aún cae en la actitud natural, porque depende de la ola de elaboraciones y suposiciones previas. La fenomenología comparte con el introspeccionismo un interés por el doblaje reflexivo como un paso clave de su enfoque al fenómeno. Pero las dos actitudes se separan. En la RF la habilidad que se moviliza se denomina colocar entre paréntesis por buenas razones, puesto que precisamente busca el efecto contrario de una introspección poco exigente: interrumpe nuestras

creencias y elaboraciones rápidas y veloces, en particular localizando y poniendo en desuso lo que creemos que "deberíamos" encontrar, o alguna descripción "esperada". Así, la RF no es una "mirada interior", sino una tolerancia respecto a la suspensión de las conclusiones que permiten desarrollar un nuevo aspecto o percepción del fenómeno. Por consiguiente, este paso no mantiene la dualidad básica sujeto-objeto, sino que desemboca en un ámbito de fenómenos en que se vuelve cada vez menos obvio cómo distinguir entre el sujeto y el objeto (esto es lo que Husserl llamó la "correlación fundamental").

Es importante reabrir el debate acerca de las diferencias clave entre el introspeccionismo (que no llevó a una sucesión fructífera) y la fenomenología (con su historia ininterrumpida). Por ejemplo, Searle quien afirma que las experiencias en primera persona son irreductibles, no hace referencia a este complejo, intelectual e histórico problema, y concluye rápidamente que la introspección es sólo otro estado mental. Por lo tanto, no se puede afirmar tener un acceso privilegiado a la experiencia y que la irreductibilidad de la experiencia "no tiene consecuencias profundas" (1992, p.118). Esta destitución del introspeccionismo y (por omisión) de la fenomenología no lleva muy lejos a Searle. A pesar de su renuncia a la introspección, eso es precisamente lo que hace en el capítulo "La estructura de la conciencia: Una introducción", que contiene doce atributos que le parecen fundamentales. ¿Con qué fundamento? ¿Realizando una introspección repentinamente válida? ¿Cómo valida estas observaciones? ¿Por qué no una lista alternativa de atributos?

La intuición no es algo erróneo

Muchas personas reaccionan con desconfianza ante la intuición. En este contexto, la capacidad intuitiva no se refiere a una fugaz inspiración quimérica. En cambio, ésta es una capacidad humana que opera constantemente en la vida cotidiana y que ha sido ampliamente tratada en los estudios de la creatividad. Piénsese en las matemáticas: en última instancia, el peso de una prueba es su naturaleza convincente, la inmediatez de la evidencia que es impuesta en nosotros, más allá de las cadenas lógicas de razonamiento simbólico. Esta es la naturaleza de

la evidencia intuitiva: nacer no de un argumento, sino del establecimiento de una claridad que es completamente convincente. Damos por sentada esta capacidad, pero hacemos poco por cultivarla de manera sistemática. Obviamente, aquí no existe contradicción entre razonamiento y deducción: la intuición sin razonamiento es ciega, pero las ideas sin intuición están vacías.

Hay vida más allá de la dualidad objetiva/subjetiva

Una de las características originales de la actitud fenomenológica es que no busca oponer lo subjetivo a lo objetivo, sino ir más allá de la escisión en su correlación fundamental. La RF nos pone rápidamente en la evidencia de que la conciencia está inseparablemente vinculada a lo que pasa más allá de ella (esto es "trascendental" en lenguaje husserliano). La conciencia no es un suceso interno privado, que tiene, al final, una existencia del mismo tipo que el mundo no-consciente externo.

Para empezar, la investigación fenomenológica no es mi "viaje privado", puesto que está destinada a otros mediante una validación intersubjetiva. En este sentido, lo que está en la actitud fenomenológica no es radicalmente diferente a los demás modos de investigación. Tal como lo señalan Hut y Shepard en su contribución a este ámbito:

> *Una analogía con la geometría de Euclides puede ser útil: una vez que especificamos la altura de los dos lados del triángulo y la medida del ángulo adyacente, se fija la altura del tercer lado y así se obtienen las medidas de los dos ángulos restantes. ¿Por qué es así? ¿En dónde radica el poder mágico del espacio? ¿Cómo puede el espacio cumplir las "leyes" de la geometría, leyes que los objetos físicos obedecen tan bien, con tan gran exactitud? (Hut y Shepard, 1996, p. 317).*

Asimismo, la evidencia empírica e intuitiva nos ha convencido de que nuestra experiencia humana, tanto la mía como la suya, sigue algún principio estructural fundamental que, como el espacio, obedece a la naturaleza de lo que nos es dado como contenidos de esta experiencia.

A través de la RF, la conciencia se presenta como un fundamento que vierte luz sobre la manera en que las nociones derivadas, como objetivo y subjetivo, pueden surgir en primer lugar. Por lo tanto, la conciencia en este tipo de examen es drásticamente diferente a la del empirismo angloamericano. No estamos interesados en una revisión privada, sino en un campo de fenómenos en que lo subjetivo y lo objetivo, así como el sujeto y otros emergen naturalmente del método aplicado y su contexto. Este es un punto que los reduccionistas y los funcionalistas olvidan a menudo. La experiencia es claramente un suceso personal, pero eso no significa que es priv*ado*, en el sentido de algún tipo de sujeto aislado que es lanzado sobre un mundo objetivo predeterminado. Uno de los descubrimientos más impresionantes del movimiento fenomenológico fue haberse dado cuenta rápidamente de que una investigación de la estructura de la experiencia humana provoca inevitablemente un cambio hacia la consideración de los muchos niveles en que mi conciencia se encuentra inextricablemente vinculada a aquellas de los demás y al mundo de fenómenos en un engranaje empático[7].

Por consiguiente, la característica oposición de los informes en primera persona versus los informes en tercera persona es engañosa. Ésta nos hace olvidar que los denominados informes objetivos en tercera persona son realizados por una comunidad concreta de personas que está encarnada en su mundo social y natural

7 E. Stein y A. Schutz fueron dos de los exploradores más activos de la empatía (*Einfhlung*) en los comienzos de la fenomenología. Para un excelente debate reciente acerca del desarrollo de la trascendencia, la empatía y la intersubjetividad en la fenomenología de Husserl ver Depraz (1996). Abrams (1996) ofrece una evocación poética de los mismos temas con una visión ecologista. Este descubrimiento fenomenológico de la empatía puede ser ilustrado para nuestros propósitos aquí, con varios estudios empíricos paralelos, es decir, algunos de sus correlatos naturales de los estudios científicos en lugar de la misma RF. A medida que los estudios acerca de la mente siguen progresando, se hace claro que como muchas otras funciones supuestamente superiores, la solidaridad y la empatía están presentes en todos los mamíferos superiores y ciertamente en los primates. Tal como de Waal (1996) argumenta convincentemente en su reciente ensayo, los monos muestran todo el espectro de inclinaciones morales y desde una temprana edad se ponen en el lugar de otro individuo, incluso aunque no existan vínculos de sangre. De esta manera, desde nuestras tempranas raíces evolutivas, el sentido de sí mismo es más adecuadamente visto como un punto holográfico que no puede ser separado de los otros puntos, múltiples y distribuidos, que constituyen nuestra inevitable ecología humana.

tanto como los informes en primera persona. Tal como B.C. Smith acertadamente pregunta: ¿Quién está en tercera? (Smith, 1996). La línea divisoria entre el rigor y la falta de él no debe ser establecida entre los informes en primera y tercera persona, sino más bien determinada por la existencia de un fundamento metodológico claro que lleve a una validación común y a un conocimiento compartido.

Se necesita una pragmática mejor

En general, mi afirmación es que la neurofenomenología es una solución natural que puede permitirnos movernos más allá del difícil problema en el estudio de la conciencia. Hay muy poco que hacer con algunos "ingredientes extra", teóricos o conceptuales, para utilizar la expresión de Chalmers. En cambio, se reconoce un campo de ignorancia *práctica* que puede ser remediado. También está claro que como todas las soluciones en la ciencia, que reformulan radicalmente un problema pendiente en lugar de tratar de resolverlo dentro de su marco original ésta tiene un potencial revolucionario, un punto al que volveré al final de este artículo. En otras palabras, en lugar de encontrar "ingredientes extra" para explicar la manera en que la conciencia emerge de la materia y el cerebro, mi propuesta reformula el problema de encontrar puentes significativos entre dos áreas fenomenales irreductibles. En este sentido específico, la neurofenomenología es una solución potencial al difícil problema que muestra de manera muy distinta lo que significa "difícil".

Tengo plena conciencia de que lo que he dicho aquí y lo que está disponible en publicaciones acerca de la reducción es limitado[8]. Esto es tanto un síntoma como una causa de la escasez relativa de trabajo reciente sobre los enfoques fenomenológicos de la mente. No se puede culpar al lector por no tener más que una leve idea de lo que quiero decir al enfatizar el gesto de la reducción, el núcleo del remedio metodológico que estoy ofreciendo aquí. Es sorprendente el hecho de que esta capacidad para tomar conciencia haya sido tan poco tomada en cuenta como una pragmática humana.

8 Vee los primeros intentos de Don Ihde (1977) para remediar esta situación.

Es como si la capacidad para el movimiento rítmico no hubiera llevado al desarrollo de la danza. Una reflexión de inspiración fenomenológica requiere estrategias para su desarrollo tal como los profesionales de las ciencias cognitivas lo han sabido por cierto tiempo (Vermesch, 1994), y como lo atestigua la tradición consciente de varias escuelas budistas (Varela *et al.)* , *1991*). Mi único comentario frente a esta pobreza relativa de la elaboración pragmática es que ésta representa un llamado urgente a la investigación para llenar esta brecha abierta. Mi propia contribución respecto a la práctica de la reducción y su formación será presentada en un próximo trabajo conjunto (Depraz *et al.* , 1996).

En occidente no hemos tenido un rico panteón de individuos con el don de la pericia fenomenológica (con notables excepciones, como Husserl o James) que hagan públicas sus investigaciones para una comunidad atenta. Por consiguiente, esta vía de investigación puede parecer ajena a muchos lectores. Sin embargo, mi opinión es que hoy, esta ausencia se encuentra en la raíz de la impenetrabilidad de la conciencia para la ciencia. Lo que se necesitan son precisamente las *estructuras de enlace* entregadas por la RF, puesto que son inmediatamente pertinentes para la experiencia (a través de su naturaleza misma) y al mismo tiempo suficientemente intersubjetivas para servir como equivalentes constructivos para el análisis externo.

3. Una circulación neurofenomenológica

Estudios de caso

En esta Sección deseo esbozar algunas áreas de la experiencia y de la vida mental con el fin de ilustrar de manera más concreta lo que podría significar, en la práctica, la circulación neurofenomenológica. De más está decir que estos estudios de caso no constituyen una prueba de lo que estoy proponiendo, ni descartan el examen detallado de otros ejemplos más interesantes para el lector. Además, en los últimos años ha existido una serie de diferentes estudios en que, aunque siguen estando bien

fundados en la tradición científica de la neurociencia cognitiva, el papel de la experiencia vivida es progresivamente más importante, a tal punto que comienza a entrar ineludiblemente en el panorama, a parte de cualquier interés en los informes en primera persona (Picton y Stuss, 1994). Es claro que, mientras más métodos sofisticados de imágenes cerebrales estén disponibles, necesitaremos sujetos con una mayor competencia en la realización de discriminaciones fenomenológicas y descripciones. Este es un problema filosófico importante, pero también es una necesidad empírica y pragmática. A continuación presentaré casos ilustrativos que tocan problemas globales y más locales.

1. Problemas globales

Puede comprenderse la atención como uno de los mecanismos básicos de la conciencia (Posner, 1994). En los últimos años, los estudios de los registros eléctricos y, más específicamente, de las imágenes funcionales del cerebro, han llevado a la identificación de redes y caminos que entregan antecedentes útiles para distinguir los sucesos cognitivos conscientes de los no-conscientes. Se pueden distinguir tres redes de atención, que toman parte en la orientación de la estimulación sensorial, la activación de patrones de la memoria y la mantención de un estado de alerta. Estos resultados indican que los mecanismos de atención son una serie distinta de procesos en el cerebro, que no están ni ubicados en unas pocas neuronas, ni son sólo el conjunto del cerebro en operación. Al mismo tiempo, es claro que las distinciones experienciales entre estas formas de atención requieren de una investigación estructural detallada de la variedad de formas en que la atención se manifiesta en la experiencia. El estudio sistemático de las estructuras y estrategias de la atención es aún, en buena parte, una tarea no cumplida. ¿Pero, cómo se hace para investigar los mecanismos neurales relacionados con la conciencia, a menos que los equivalentes experienciales puedan ser suficientemente discriminados, reconocidos y enseñados?

Conciencia del tiempo presente. La temporalidad es inseparable de toda experiencia y, en varios horizontes de duración, del ahora

presente a una duración total. Un nivel de estudio es precisamente la experiencia del tiempo inmediato, la estructura de un ahora como tal o como en la acertada frase de James (1912) "el presente engañoso". Éste ha sido un tema tradicional en los estudios fenomenológicos, que describe una estructura arbórea del presente de tres partes con sus ramas constitutivas en el horizonte del pasado y del futuro, las llamadas pretensiones y retenciones (Husserl, 1966; MacInerny, 1991). De hecho, estas invariantes estructurales no son compatibles con la representación de punto continuo del tiempo lineal que hemos heredado de la física. Sin embargo, éstas se vinculan naturalmente a un conjunto de conclusiones de la neurociencia cognitiva que dicen que existe un tiempo mínimo requerido para la emergencia de los sucesos neurales que se correlacionan con un suceso cognitivo (Dennett y Kinsbourne, 1992). Este marco temporal no compresible puede ser analizado como una manifestación de la integración neuronal a largo plazo en el cerebro vinculada con una sincronía generalizada (Singer, 1993; Varela, 1995). Este vínculo aclara la naturaleza de las invariantes fenomenológicas a través de una reconstrucción dinámica que subyace a ellas, además de dar al proceso de sincronía un contenido experiencial tangible. He desarrollado este caso de circulación neurofenomenológica con más detalle en otro trabajo (Varela, 1996).

Imagen del cuerpo y movimiento voluntario. La naturaleza de la voluntad, tal como se expresa en el inicio de una acción voluntaria, es inseparable de la conciencia y de su examen. Estudios recientes otorgan un papel importante a los correlatos neurales, que preceden y preparan la acción voluntaria y al papel de la imaginación en la constitución de un acto voluntario (Libet, 1985; Jeannerod, 1994). Aunque la acción voluntaria es sobre todo una experiencia vivida, de la que se ha hablado mucho en la literatura fenomenológica, más específicamente en el papel de la encarnación como cuerpo vivido (*corps propre*, Merleau-Ponty, 1945), y además en la estrecha relación entre el cuerpo vivido y su mundo (*Leibhaftigkeit*). Por ejemplo, el dolor es una *"qualia"* interesante que revela esta dimensión de encarnación de manera más viva, y este estudio fenomenológico arroja sorprendentes percepciones de la imagen del

cuerpo y de su relación con los correlatos neurofisiológicos (Leder, 1991). Una vez más, es esencial un análisis fenomenológico de la acción voluntaria y de la encarnación, pero, hasta cierto punto, sólo parcialmente desarrollado.

2. *Problemas locales*

La completitud perceptiva (*filling-in*) como se utiliza en la ciencia visual, consiste en completar espontáneamente un percepto de manera que la apariencia (a saber, un contorno visual) es distinta del correlato físico (a saber, límites discontinuos, como en el caso de los populares contornos ilusorios). Estos problemas pueden ser estudiados incluso a nivel celular y suscitan más preguntas con relación a la distinción experiencial de las apariencias. De hecho, la información neuronal acerca del suplemento parece correlacionarse bien con la conclusión que la RF había sacado hacía algún tiempo: existe una diferencia importante entre "ver cómo", la apariencia visual y "ver qué", un juicio visual (Pessoa *et al.*, 1996). Esta conclusión se opone a la de Dennett (1991), para quien la conciencia es "decir y no mostrar". Estos son problemas que sólo pueden ser resueltos con la convergencia concertada de informes externos y directos.

Periferia y centro. Es interesante para nosotros que una serie de estudios haya vuelto a considerar algunos problemas fenomenológicos tradicionales, como la estructura de dos partes del ámbito de la conciencia entre un centro y una periferia. Esto proviene en gran parte de la influencia de William James, pero llevado a los modernos protocolos de laboratorio. En estos estudios, la experiencia crucial para estudiar e identificar para su mejora, es el sentimiento de "corrección", que aquí se presenta como un resumen de la integración cognitiva, que representa el grado de armonía existente entre el contenido consciente y su trasfondo paralelo inconsciente (Mangan, 1993).

Emoción. En los últimos años, se han visto importantes avances en la comprensión de los correlatos cerebrales de las emociones; la

separación entre el razonamiento y las emociones está desapareciendo rápidamente (Damasio, 1994; Davidson y Sutton, 1994). La evidencia recalca la importancia de las estructuras específicas tales como la amígdala, la lateralización de los procesos y en el papel del despertar sexual en la memoria emocional. Sin embargo, estos estudios se basan por completo en protocolos verbales y los problemas de la competencia para la distinción emocional y los patrones de relaciones entre humor, emoción y razones necesitan ser dirigidos explícitamente en esta fase de la investigación.

La evocación de estos estudios de caso intenta otorgar un trasfondo concreto para tratar más a fondo la preocupación central del programa neurofenomenológico que presento aquí. Por una parte, tenemos un proceso de emergencia con atributos neurobiológicos bien definidos. Por otra, una descripción fenomenológica que se vincula directamente con nuestra experiencia vivida. Para realizar mayores progresos necesitamos análisis y técnicas vanguardistas de parte del lado científico y un desarrollo muy constante de la investigación fenomenológica para los objetivos de la investigación misma.

¿Espero que la lista de invariantes estructurales relacionadas con la experiencia humana aumente *ad infinitum*? No, por cierto. Supongo que se puede esperar que el horizonte de temas fundamentales converja en un corpus de conocimiento bien integrado. La rapidez con que esto suceda y cuándo, dependerá , por supuesto, del ritmo en el que se constituya una comunidad de investigadores comprometida con esta modalidad de investigación y cree más estándares de evidencia.

La hipótesis de trabajo

Esto me lleva a mi punto de partida: sólo un informe equilibrado y disciplinado del lado externo y experiencial de un problema puede hacernos estar un paso más cerca de salvar la distancia entre la mente biológica y la mente experiencial. Permítanme ser más explícito sobre mi hipótesis de trabajo básica para una "circulación" entre el análisis externo y fenomenológico:

La Hipótesis de Trabajo de la Neurofenomenología

Las referencias fenomenológicos acerca de la estructura de la experiencia y sus equivalentes en la ciencia cognitiva se relacionan unas con otras a través de restriccionesrecíprocas.

El punto clave aquí, es que a través del énfasis de una co-determinación de ambos informes se pueden estudiar los puentes, desafíos, percepciones y contradicciones que existen entre ellos. Esto significa que ambas áreas de fenómenos tienen el mismo estatus para exigir una atención total y respeto por su especificidad. Es bastante fácil ver cómo los informes científicos aclaran la experiencia mental, pero es característico que se ignore la dirección recíproca de la experiencia hacia la ciencia. ¿Qué entregan los informes fenomenológicos? Al menos dos aspectos principales del panorama global. En primer lugar, sin ellos el carácter directo de la experiencia desaparece, o se transforma en un misterio. En segundo lugar, los informes estructurales entregan limitaciones acerca de las observaciones empíricas.

El estudio de la experiencia no es una parada conveniente en nuestro camino hacia una explicación real, sino un participante activo en su derecho propio. Claramente, en este programa de investigación, como en todos los demás que merecen este nombre, se acumula con lentitud un conjunto de evidencia, mientras que otros aspectos son más oscuros y difíciles de aprovechar. Es evidente que los estudios de caso nombrados con anterioridad necesitan substancialmente mayor desarrollo, pero espero que esté claro cómo estos comienzan a entregar una perspectiva "estereoscópica" acerca de los muchos problemas globales y locales en que la experiencia y la ciencia cognitiva se vuelven socios activos.

Esta demanda de una circulación disciplinada es un estándar más preciso y más exigente que el "equilibrio reflexivo" propuesto por Flanagan (1992) o la "proyección consciente" ampliada por Velmans (1996). Aunque existe similitud de intención con lo que estoy proponiendo, ellos no proponen fundamentos metodológicos

explícitos o nuevos para llevar a cabo estas intenciones. Seguramente, ésta es una mejora de Searle, quien insiste en el hecho de que toma una actitud naturalista y en que la conciencia es "obviamente" una emergencia. Sin embargo, su naturalismo no funciona en su libro: no existe ni una sola línea acerca de los mecanismos explícitos y, por lo tanto, su naturalismo es estéril. En fin, la hipótesis presentada aquí entrega un camino explícito para conducir la investigación en la ciencia cognitiva como si importara tanto la fisiología cerebral como la experiencia mental. Así, por ejemplo, un mecanismo de integración a gran escala en el cerebro, como la sincronía neural en la banda gama, también debería ser validada sobre la base de su capacidad para llegar a comprender los informes en primera persona de contenidos mentales, como la duración. Las preguntas empíricas deben ser guiadas mediante evidencia en primera persona. Esta limitación doble no se aplicaría a las descripciones que no se relacionan directamente con el nivel de experiencia, por ejemplo, para las respuestas celulares o la difusión de neurotransmisores.

La afirmación acerca de los niveles apropiados de descripción entre los sucesos cerebrales y la conducta no es, por supuesto, nueva y controversial excepto para aquellos que son extremadamente reduccionistas. La novedad de mi propuesta, consiste en que estos informes d*isciplinados en primera persona* deberían ser un elemento integral de la validación de una propuesta neurobiológica y no sólo una información casual o heurística. Esta es la razón por la cual elegí describir la situación mediante la hipótesis de que ambos informes tienen *limitaciones* mutuas uno sobre otro.

Aun así, ¿ésta no es sólo una versión aumentada de la conocida teoría de la identidad (o al menos un homomorfismo) entre la experiencia y los informes neurocientíficos cognitivos? Realmente no, puesto que estoy afirmando que los correlatos deben ser establecidos, no sólo como una cuestión de compromiso filosófico o supuesto físico, sino como un examen metodológico sólido de las invariantes experienciales. Una vez más, ésta es una cuestión de pragmática y aprendizaje de un método, no de un argumento *a priori* o de una completud teórica.

En cambio, una tesis de psico-identidad más convencional estudia una forma de razonamiento que Pessoa *et al.* (1996) llama proposiciones vinculantes (de acuerdo al D. Teller). Estas son proposiciones del tipo:

Φ se parece a Ψ=> Φ explica Ψ

en que Φ representa los términos neuro-psicológicos y Ψ representa los términos fenoménicos, y el operador de implicancia tiene un sentido condicional: si los sucesos empíricos "se parecen" a los sucesos fenoménicos, entonces estos son explicados. Un excelente ejemplo es el entusiasmo de Crick, cuando habla de los correlatos neuronales únicos asociados con el súbito cambio de la experiencia en las figuras visuales binoculares rivales (Leopold y Logothetis, 1996), que él asimila como una explicación de esta forma de conciencia visual. Estos tipos de puentes no son satisfactorios porque dejan el problema intacto. Aún tenemos que lidiar con la naturaleza del problema: cómo se relacionan estas unidades neuronales con el resto de la actividad cerebral, cómo adquieren su sentido y, especialmente, qué, en ellas, las transforma en un suceso experiencial. Volvemos a empezar desde cero con el difícil problema intacto. La diferencia en la estrategia de investigación propuesta por la neurofenomenología radica en que estos puentes no son del tipo "se parece", sino que están construidos mediante limitaciones mutuas y son validados a partir de ambas áreas fenomenológicas, en que los términos fenoménicos se presentan como términos explícitos *directamente* relacionados con la experiencia a través de un examen riguroso (por ejemplo, reducción, invarianza y comunicación intersubjetiva).

Esta hipótesis de trabajo tiene algunos puntos de similitud con la noción de "coherencia estructural", presentada por Chalmers entre sus tres principios básicos para la estructura de la conciencia. Además, "precisamente porque las propiedades estructurales de la experiencia son accesibles e informables, esas propiedades estarán directamente representadas en la estructura de la conciencia " (1995, p. 213). Esto es bastante correcto según mi punto de vista, pero es fatalmente incompleto tal como se plantea, al menos en relación a

dos problemas clave planteados aquí. En primer lugar, esta estructura de experiencia necesita un método para el estudio y la validación, y no es suficiente con afirmar simplemente que podemos trabajar con la estructura de la conciencia. En segundo lugar, no existe valor ontológico en el principio de Chalmers, puesto que él supone que la conciencia es un término ontológico adicional. En nuestro caso, la experiencia fenoménica representa un nivel ontológico irreductible, pero conserva su cualidad de inmediatez porque cumple un papel en la coherencia estructural a través de sus contenidos intuitivos y, por lo tanto, mantiene viva su conexión directa con la experiencia humana, en lugar de presionarla a la abstracción.

Esto es todo lo que importa: se obtiene un informe intelectualmente coherente de la mente y la conciencia en que el polo experiencial entra directamente en la formulación de la exploración completa, haciendo referencia directa a la naturaleza de nuestra experiencia vivida. La "dificultad" y el misterio favorecen un programa de investigación abierto a su estudio incondicional, en que la estructura de la experiencia humana juega un papel central en nuestra explicación científica. En todos los estudios funcionalistas lo que falta no es la naturaleza coherente de la explicación, sino su distanciamiento de la vida humana. Sólo la incorporación de la vida humana borrará esta falta; no algún "ingrediente extra" ni una "posición teórica" profunda[9].

Del mismo modo, habría que esperar del enfoque neurofenomenológico una percepción completamente nueva de los

9 Casualmente, la propia "solución" teórica (o ingrediente extra) de Chalmers es su noción de "doble información" derivada de la antigua teoría shannoniana de los símbolos, incorrectamente conocida como una "teoría de la información". En un libro, en que se muestra esta clara agudeza intelectual, me quedé atónito al ver que al final Chambers argumenta que la mejor elección es revivir una tradición cibernética, en gran parte transformada después de su inicio en herramientas verdaderamente informativas a través del trabajo realizado en enfoques computacionalistas, conexionistas y de encarnación de la cognición, nunca tratados por Chalmers en este contexto. Incluso, asumiendo la postura de que se necesita un "ingrediente extra", simplemente no veo lo que podría posiblemente derivar de este paso y ni tampoco lo hacen los científicos que han hecho comentarios al respecto. Véase, por ejemplo, la resencisión más bien positiva de Koch (Koch, 1996).

mecanismos empíricos. ("Así, ¿qué agregaría a la ciencia cognitiva con su método que nosotros ya no sepamos?"). Seguramente, el enfoque de la RF proporciona ideas interesantes en lo que respecta a la estructura de la vida mental (cf. los casos de temporalidad o sustitución), pero su fuerza principal radica en que lo hace de una manera que hace nuestra experiencia reconocible. Estoy totalmente seguro de que este inesperado giro de segundo orden será el más difícil de apreciar para aquellos investigadores de constante tendencia funcionalista.

4. Conclusión

La conciencia: ¿un difícil problema o una bomba de tiempo?

Prácticamente desde su inicio, la ciencia cognitiva ha estado comprometida con una serie muy explícita de ideas clave y metáforas, que pueden denominarse *representacionalismo*, para el cual la distinción interior-exterior es la pieza central: un exterior (un mundo lleno de características) representado interiormente a través de la acción de dispositivos perceptuales complejos. En los últimos años, ha habido un cambio lento pero seguro hacia una orientación alternativa, a la que he contribuido y defendido durante muchos años (ver Varela, 1979; Varela *et al.*, 1991). Esta orientación difiere del representacionalismo, por tratar a la mente y al mundo como mutuamente superpuestos, de ahí los términos calificadores *encarnado*, situado o ciencia cognitiva *enactiva*.

No puedo elaborar aquí la situación actual de la ciencia cognitiva, pero mi presente propuesta respecto al estudio de la conciencia se alinea con esas mayores preocupaciones. Parece inevitable tomar la tendencia hacia la encarnación como un paso adelante en dirección de una consideración de principios de la *encarnación como experiencia vivida*. En nuestro libro (Varela et al., 1991), primero destacamos la circularidad intrínseca en la ciencia cognitiva en la cual el estudio de los fenómenos mentales siempre es el de una persona que experimenta. Afirmamos que la ciencia cognitiva no puede escapar a esta circulación, y en

lugar de eso, debe cultivarla. Nosotros nos inspiramos explícitamente en las tradiciones asiáticas, en particular en el budismo, como manifestaciones vivientes de una fenomenología activa y disciplinada. La intención de ese libro no era detenernos demasiado en las tradiciones asiáticas *per se*, sino utilizarlas como un reflejo lejano de lo que necesitamos cultivar en nuestra ciencia y en la tradición occidental.

La presente proposición avanza un paso más con relación a lo que fue tomado en el libro, concentrándose en el problema clave de la metodología. Espero haber tentado al lector a pensar que tenemos frente a nosotros la posibilidad de una búsqueda abierta de pasajes resonantes entre la experiencia humana y la ciencia cognitiva. Sin embargo, el precio es tomar seriamente los informes en primera persona como un ámbito de fenómenos válidos. Y más allá de eso, crear una *tradición* continua de examen fenomenológico que hoy, en general, casi no existe en nuestra cultura y ciencia occidental.

Se debe tomar en serio el doble desafío que mi propuesta representa. En primer lugar, exige un re-aprendizaje y una maestría de la habilidad de descripción fenomenológica. No hay motivo para que esto deba ser diferente de la adquisición de cualquier *know-how*, como aprender a tocar un instrumento o hablar un nuevo idioma. Cualquiera que se dedique al aprendizaje, ya sea en música, lenguaje o pensamiento, provocará un cambio en la vida cotidiana. Esto es lo que figura como artículo cuarto en la RF: el aprendizaje constante y disciplinado conlleva una transformación y así cualquier otra cosa que hagamos de manera constante. Esto está bien, si rechazamos la suposición (como yo lo hago) de que existe un tipo de estándar bien definido para lo que debería ser considerado como experiencia real o normal: la experiencia parece ser inherentemente abierta y flexible y, por lo tanto, no es una contradicción decir que la formación constante en un método puede hacer disponibles aspectos de la experiencia que no lo eran antes. El objetivo de la RF es, entre otras cosas, superar la costumbre de la introspección automática; no necesitamos llevar

luto por lo que se puede perder, sino un interés en lo que puede aprenderse[10].

El segundo desafío que mi propuesta representa, es el de un llamado a transformar el estilo y los valores de la comunidad investigadora misma. No podemos esperar avanzar y romper el ciclo histórico de rechazo-fascinación de la conciencia en la filosofía de la mente y la ciencia cognitiva, a menos que aceptemos que, en este momento de la historia intelectual y científica, se necesita algún re-aprendizaje radical. Mi propuesta implica que cada buen alumno de ciencia cognitiva, que también esté interesado en problemas a nivel de la experiencia mental, debe inevitablemente alcanzar un nivel de maestría en el examen fenomenológico con el fin de trabajar seriamente con los informes en primera persona. Pero esto sólo sucederá cuando toda la comunidad se adapte con un correspondiente cambio de actitud en relación con los tipos de argumento aceptables, la evaluación de estándares y las políticas editoriales en las principales publicaciones científicas de manera que esta competencia adicional se transforme en una dimensión importante para un investigador joven. Para la antigua tradición de ciencia objetivista esto suena a anatema, y lo es. Sin embargo, no es una traición a la ciencia: es un complemento y una extensión necesarios. La ciencia y la experiencia se limitan y se modifican la una a la otra como en una danza. Aquí radica el potencial de transformación y es también la clave de las dificultades que esta posición ha encontrado dentro de la comunidad científica. Esto requiere que dejemos atrás cierta imagen de cómo se hace ciencia y poner en duda un estilo de formación científica que es parte de la misma estructura de nuestra identidad cultural.

---

10 Dreyfus (1993) en una reseña sobre nuestro libro nos criticaba por dar énfasis a la transformación que acompaña al aprendizaje de la observación fenomenológica, puesto que ella misma interfiere con la "experiencia diaria". Sería un error, si uno creyera que expone una "capa más profunda" al adquirir una habilidad, como, por ejemplo, la RF o el dedicarse a una práctica como la conciencia atenta, lo que no era en absoluto nuestra afirmación. Incluso Dreyfus tendría que llegar a la conclusión de que no hay una posición privilegiada para decirnos qué es lo que vale como experiencia "real". Claramente, él ha malinterpretado lo principal: la reducción fenomenológica no "revela" un fundamento objetivo, sino que da a luz nuevos fenómenos dentro del campo experiencial, en un abanico de posibilidades múltiples.

En resumen: ¿Cuál es la idea?

Permítanme concluir resumiendo los puntos principales que he planteado en esta reacción frente al "difícil" problema de la conciencia, basado en una propuesta explícita para su solución.

El argumento:

- Al igual que la idea básica de Chalmers, creo que la experiencia directa, vivida, es un *campo de fenómenos* auténtico, irreductible a cualquier otra cosa. Mi afirmación es que no existe una posición teórica o un ingrediente "extra" esencial que posiblemente pueda salvar esta brecha.
- Ahora bien, este campo de fenómenos requiere un méto*do* apropiado y riguroso y una pragmática para su estudio y análisis.
- La orientación de este método se inspira en el estilo de investigación de la *fenomenología* con el fin de constituir una amplia *comunidad de investigación* y un programa de investigación.
- Este programa de investigación busca articulaciones mediante limitaciones mutuas entre el campo de los fenómenos revelado por la experiencia y el campo correlativo de fenómenos establecido por las ciencias cognitivas. He denominado *neurofenomenología* a este punto de vista.

Las consecuencias:

- Sin una expansión radical del estilo de trabajo de la tradición científica y el establecimiento de un programa de investigación aproximadamente en esta línea, el enigma del lugar de la experiencia en la ciencia y en el mundo seguirá retornando, ya sea para ser explicado convincentemente o bien para reafirmar que es demasiado difícil, dado lo que sabemos.
- La naturaleza "difícil" se reformula en dos sentidos:

  (1) es difícil enseñar y estabilizar nuevos métodos para explorar la experiencia,

(2) es difícil cambiar los hábitos de la ciencia con el fin de que ésta acepte que necesita nuevas herramientas para la transformación de lo que significa conducir una investigación acerca de la mente y para la formación de las generaciones futuras.

REFERENCIAS BIBLIOGRÁFICAS

Abrams, D. *(1996), The Spell of the Sensous: Perception and Language in a More Than Human World (*Nueva York: Pantheon).

Baars, B. (1988), A *Cognitive Theory of Consciousness* (Cambridge: Cambridge University Press).

R. Bernet, (1994), *La Vie du Sujet* (París: Presses Universitaires de France).

Calvin W. (1990), *Cerebral Symphony: Seashore Reflections on the Structure of Consciousness* (Nueva York: Bantam Books).

Chalmers, D.J. (1995), 'Facing up to the problem of consciousness', *Journal of Consciousness Studies,* 2 (3), pp. 200-19.

Chalmers, D.J. (1996), *The Conscious Mind: In Search of a Fundamental Theory (*Nueva York: Oxford University Press).

Churchland, P.S. and Sejnowski, T. (1992), *The Computational Brain* (Cambridge, MA: MIT Press). Crick, F. (1994), *The Astonishing Hypothesis* (Nueva York: Scribners).

Damasio, A. (1994), *Descartes' Error: Emotion, Reason and the Human Brain* (Nueva York: Grosset/Putnam).

Davidson, R. and Sutton, S. (1994), 'Affective neuroscience: the emergence of a discipline', *Curr. Opinion Neurobiol.,* 5, pp. 217-24.

de Waal, F. (1996), *Good Natured: The Origins of Right and Wrong in Humans and Other Animals* (Cambridge: Harvard University Press).

Dennett, D. (1991), *Consciousness Explained* (Boston: Little Brown).

Dennett, D. and M. Kinsbourne (1992), 'Time and the observer: the where and when of time in the brain', *Behavioral and Brain Sciences,* 15, pp. 183-247.

Depras, N. (1996), I*ncarnation et Transcendence* (París: J.Vrin).

Depraz, N., Varela, F. and Veermersch, P. (1996), *Exploring Experience with a Method* (forthcoming).

Dreyfus, H. (ed. 1982), *Husserl: Intentionality and Cognitive Science* (Cambridge, MA: MIT Press).

Dreyfus, H. (1993), Review *of*'The Embodied Mind', *Mind,* **102,** pp. 542.
Edelman, G. (1989), *The Remembered Present: A Biological Theory of Consciousness* (Nueva York: Basic Books).
Flanagan 0. (1992), *Consciousness Reconsidered* (Cambridge, MA: MIT Press).
Globus, U. (1995), *The Post-Modern Brain* (New York: Benjamin).
Howe, R.B. (1991), Introspection: A reassesment, *New Ideas in Psychology,* 9, pp. 244.
Husserl, H. (1962), *The Idea of Phenomenology* (The Hague: M. Nijhoff).
Husserl, E. (1966), *Zür Phänomenologie des Inneren Zeitbewusstseins (1893-1917),* ed. R. Bohm (The Hague: M.Nijhoff). Partial English translation: *The Phenomenology of Internal Time Consciousness (*Bloomington: Indiana University Press, 1996).
Hut, P. and Shepard, R. (1996), 'Turning the hard problem upside down and sideways', *Journal of Consciousness Studies,* 3 (4), pp. 313-29.
lhde, D. (1977), *Experimental Phenomenology* (Nueva York: Open Court).
Jackendoff, R. (1987), *Consciousness and the Computational Mind* (Cambridge, MA: MIT Press). James, W. (1912/1995), *The Principles of Psychology* (Cambridge: Harvard University Press).
Jeannerod, R. (I 994), 'The representing brain: neural correlates of motor intention and imagery', *Behavioral and Brain Sciences,* 17, pp. 187-245.
Johnson, M. (1987), *The Body in the Mind: The Bodily Basis of Imagination, Reason and Feeling* (Chicago: Chicago University Press).
Klein, P. and Wescott, M.R. (1994), 'The changing character of phenomenological psychology', Cana*dian Psychology,* 35, pp. 133-57.
Koch, C. (1996), Review of 'The Conscious Mind', *Nature, 381,* p. 123.
Langsdorf, L. (1985), Review of Dreyfus (1982), *Husserl Studies,* 3, pp. 303-11.
Leder, D. (1991), *The Dissapearance of the Body* (Chicago: Chicago University Press).
Libet, B. (1985), 'Unconscious cerebral initiative and the role of conscious will in voluntary action', *Behavioral and Brain Sciences,* 8, pp. 5296.
Leopold, D. and Logothetis, N. (1996), 'Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry', *Nature,* 379, pp. 549-53.
Lyons, W. (1986), *The Disappearance of lnstrospection* (Cambridge, MA: MIT Press).

Lyotard, J-F. (1954), *La Phénoménologie* (París: Presses Univ. de France).
McGinn, C. (1991), *The Problem of Consciousness* (Oxford: Blackwell).
Mclnerney, P. (1991), *Time and Experience* (Philadelphia: Temple University Press).
McIntyre, R. (1986), 'Husserl and the representational theory of mind', *Topoi,* 5, pp. 101-13.
Mangan, B. (1993), 'Taking phenomenology seriously: the «fringe» and its implications for cognitive research', *Consciousness and Cognition, 2,* pp. 89-108.
Mazoyer, B. Ronald and Fox (ed. 1995), International Congress on the Functional Mapping of the Human Brain, *Human Brain Map.* Suppl. 1.
Merleau-Ponty, M. (1945), *La Phénoménologie de la Perception* (París: Gallimard).
Nagel, T. (1986), *The View from Nowhere* (New York: Oxford University Press).
Pessoa, L., Thompson, E and Noë, A. (1996), 'Finding out about filling in', B*ehavioral and Brain Sciences* (in press).
Petitot, J. (ed. 1995), 'Sciences cognitives et phénoménologie', *Archives de Philosophie,* 58 (4), pp. 529-631.
Petitot, J., Roy, J. M., Pachoud, B. and Varela, F. (ed. 1996), *NaturalizingPhenomenology: Contemporary Issues in Phenomenology and Cognitive Science* (Stanford: Stanford University Press). In press.
Picton,T. and Stuss, D. (1994), 'Neurobiology of conscious experience', *Current Biology,* 4, pp. 256-65.
Posner, M.I. (1994), 'Attention: the mechanisms of consciousness', *Proc. Natl. Acad. Science (USA),* 91, pp. 7398-4003.
Posner, M & Raichle, M (1992), *Images of the Mind* (Nueva York: Scientific American Library).
Roy, J.M. (1995), 'Le «Dreyfus bridge»: Husserlianism and Fodorism', *Archives de Philosophie,* 58, pp. 5339.
Searle, J. (1992), *The Rediscovery of the Mind* (Cambridge, MA: MIT Press)
Smith, B.C. (1996), 'Who's on third? Subjectivity at the physical basis of consciousness', Cons*ciousness Research Abstracts, p.* 53.
Singer, W. (1993), 'Synchronizafion of cortical activity and its putative role in information processing and learning', *Ann. Rev. Physiol.,* 55, pp. 349-74.
Spiegelberg, F. (1962), *The Phenomenological Movement,* 2 vols. 2ed. (The Hague: Martinus Nihjoff).

Thompson, E. and F. Varela (1996), *Why the Mind is not in the Head* (Cambridge: Harvard University Press, forthcoming).
Varela, F. (1979), *Principles of Biological Autonomy* (Nueva York: North-Holland).
Varela, F. (1995), 'Resonant cell assemblies: A new approach to cognitive functioning and neuronal synchrony', B*iol. Res.*, **28**, pp. 81-95.
Varela, F. (1996), 'The specious present: A neurophenomenology of nowness', in Petitot *et al.* (1996).
Varela, F. J., Thompson, E. and Rosch, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience* (Cambridge, MA: MIT Press).
Velmans, M. (1996), The *Science of Consciousness* (Londres: Routledge).
Vermersch, P. (1994), L*'Entretien d'Explicitation* (París: ESF Editeur).

METODOLOGÍAS EN PRIMERA PERSONA

Qué, por qué, cómo

Este texto es la introducción de un libro enteramente dedicado a una visión de conjunto de los métodos disponibles para acceder a descripciones en primera persona. Se trata de una visión sintética y que apunta a los problemas y las controversias, más que a proponer un punto de vista, como es el caso en el texto sobre la neurofenomenología. El hecho de ser una introducción a un libro explica también su estilo de redacción.

1: Adentro-afuera: una división engañosa

Cuando hablamos de un acontecimiento en primera persona, queremos decir una *experiencia* vivida, asociada con eventos cognocitivos y mentales. A veces se utilizan términos como "conciencia fenoménica" e incluso "*qualia*", pero es natural hablar de "experiencia consciente" o simplemente "experiencia". Los términos utilizados aquí implican que los procesos estudiados (visión, dolor, memoria, imaginación, etc.) son relevantes *para* y se manifiestan *a* un "sí mismo" o "sujeto" que puede dar cuenta de ellos: es decir estas experiencias tienen un lado "subjetivo".

Por el contrario, las descripciones en tercera persona tienen que ver con las experiencias descriptivas asociadas al estudio de otros fenómenos naturales. Aun cuando en la ciencia encontremos siempre

(F.J.Varela, First-person methodologies: why, when and how, in: F. Varela and J.Shear (Eds.), *The View from Within: First-Person Methodologies in the Study of Consciousness*, Imprint Academic, 1999.)

un agente humano que produce y entrega las descripciones, los contenidos de éstas (por ejemplo de reacciones bioquímicas, agujeros negros o voltajes sinápticos) no están asociados de manera clara o inmediata con tal agente. Las características que las definen se refieren a propiedades de los eventos que no se manifiestan directamente en la esfera mental-experimental del agente y sólo pueden ligarse a dicha esfera de manera indirecta (a través de la vida en el laboratorio, los distintos modos de comunicación científica, etc.). Estas descripciones "objetivas" tienen por cierto una dimensión subjetiva-social, pero esta queda oculta en la práctica social de la ciencia. La referencia evidente y directa es el "objetivo", el "exterior", el contenido de la ciencia referida a diversos fenómenos naturales, como la física y la biología.

Ahora bien, los desarrollos más recientes de la filosofía y la historia de la ciencia sugieren que esta aparente objetividad no puede caracterizarse por una relación con cosas-allá-afuera de manera independiente de los contenidos mentales-aquí-adentro. La ciencia está permeada por las regulaciones sociales y de procedimiento conocidas como método científico, que permite la constitución de un corpus de conocimiento compartido acerca de los objetos naturales. El eje de esta constitución es la verificación pública y su validación mediante complejos intercambios humanos. Lo que consideramos objetivo es aquello que podemos traspasar de la esfera individual a un cuerpo de conocimiento regulado. Este cuerpo de conocimiento es, inevitablemente, en parte subjetivo por cuanto depende de la observación y la experiencia individual, y en parte objetivo ,por cuanto está limitado y regulado por los fenómenos naturales, empíricos.

Este breve recordatorio de que lo subjetivo está siempre implícito en lo objetivo nos lleva necesariamente a revisar detenidamente la distinción entre lo objetivo y lo subjetivo como una absoluta demarcación entre el afuera y el adentro. *Mutatis mutandis*, lidiar con fenómenos subjetivos no es lo mismo que lidiar, como se asume generalmente, con experiencias puramente privadas. Lo subjetivo está intrínsecamente abierto a la validación intersubjetiva, si somos capaces de proveernos de un método y procedimiento que así lo permita. Uno de los propósitos centrales de los estudios

sobre las metodologías en primera persona es justamente examinar algunas de las aproximaciones más importantes y recientes orientadas a entregar las bases de una *ciencia de la conciencia que incluya a la primera persona, la experiencia subjetiva como un componente explícito y activo.*

2. Primera y tercera persona: la necesaria circulación

Expuesto el problema de esta manera, la pregunta que surge naturalmente es aquella referida al status de la primera persona. En un sentido elemental, la respuesta no puede ser dada *a priori* y sólo puede provenir de la exploración de este campo fenoménico, tal como hacen las contribuciones presentadas aquí. Sea como sea, expongamos en primer lugar algunos temas delicados, para evitar algunos malentendidos recurrentes.

Primero, explorar los relatos en primera persona no es lo mismo que afirmar que estos relatos tienen un acceso privilegiado a la experiencia. No es necesario suponer aquí nada incorregible, final, fácil o apodíctico sobre los fenómenos subjetivos, y hacerlo de otra manera sería confundir el carácter inmediato de los fenómenos subjetivos con su modo de *constitución* y *evaluación*. Se habría ahorrado mucha tinta con sólo distinguir la irreductibilidad de las descripciones en primera persona de su status epistémico.

Segundo, un punto crucial en estos estudios en primera persona ha sido subrayar la necesidad de superar la actitud de "simplemente-echar-un-vistazo" en relación con la experiencia. La aparente familiaridad que tenemos con la vida subjetiva debe dar paso al examen cuidadoso de aquello a lo que podemos acceder y de aquello a lo que no podemos acceder, y a cómo esta diferencia no es rígida sino variable. Es aquí donde la metodología se hace crucial: sin un examen sistemático no podemos producir descripciones fenoménicas que sean lo suficientemente ricas y sutilmente interconectadas como para distinguirlas de aquellas hechas en tercera persona. La pregunta central es: ¿cómo *hacerlo* realmente? ¿Hay alguna evidencia de que se *puede* hacer? Y en caso positivo ¿con qué *resultados*?

Tercero, sería fútil tratar aisladamente las descripciones en primera persona. Es necesario armonizarlas y limitarlas mediante la construcción de buenos *vínculos* con los estudios en tercera persona. En otras palabras, no nos interesa generar un nuevo debate sobre la controversia filosófica en torno a la división entre la primera y tercera persona (sobre la cual existe, por lo demás, una gran cantidad de literatura). Para hacer posible lo que nos proponemos, buscamos metodologías que puedan entregar un claro vínculo con descripciones objetivas, empíricas. Esto a menudo implica una mediación intermedia, una posición de segunda persona, tal como discutiremos más adelante. El resultado de este estudio debería llevarnos a una perspectiva integral o global donde ni la experiencia ni los mecanismos internos tengan la última palabra. La perspectiva global requiere por tanto establecer de modo explícito *restricciones mutuas,* una determinación e influencia recíprocas (Varela, 1996). En pocas palabras, nuestra posición sobre las metodologías de la primera persona es la siguiente: no dejarlas nunca de lado pero no olvidar tampoco las referencias de tercera persona.

Esta aproximación *pragmática* y bien aterrizada da el tono de este trabajo. En términos generales, lo que surge de esta reflexión es que, a pesar de todas las ideas preconcebidas, repetidas de manera irreflexiva en la literatura reciente sobre la filosofía de la mente y la ciencia cognitiva, los métodos de primera persona están a nuestro alcance y *pueden* constituier un aporte valioso a la ciencia de la conciencia. esto no se funda en argumentos a priori sino en ejemplos específicos de conocimiento práctico, en estudios de casos.

3. La noción de dato fenoménico

La noción de dato fenoménico puede proporcionar un elemento de base común para los planteamientos de primera persona y tercera persona, en relación a la pregunta recién planteada, de modo que merece una breve discusión, dado su papel clave en los estudios en primera persona. Gran parte de la literatura moderna se refiere al "vacío explicativo" entre las inteligencias computacional

y fenomenológica[1], para usar la terminología de Jackendoff (1987). Como ya hemos señalado, la segunda aparece generalmente referida en términos de subjetividad, conciencia o experiencia. Es importante examinar estos diferentes conceptos y ver cómo están relacionados con el más básico de *datos fenoménicos* (Roy *et. al.*, 1998).

A pesar de la variedad terminológica en uso, parece haber consenso en que la expresión de Thomas Nagel "qué significa ser" resume bien lo que aquí está en juego. Claramente "lo que significa ser" un murciélago o una persona se refiere a cómo las cosas (todas las cosas) se ven cuando uno es un murciélago o una persona. En otras palabras, ésta es simplemente otra manera de hablar de aquello que los filósofos, desde los presocráticos en adelante, han llamado fenomenalidad. Un fenómeno, en el sentido más original de la palabra, es una aparición y por tanto algo relacional. Es lo que algo es para algo distinto: es un *ser para* en oposición al ser en sí mismo independientemente de su aprehensión por otra entidad.

La fenomenalidad ciertamente es un factor crucial en el ámbito de los seres vivos. Por ejemplo, es evidente que un organismo con un sistema de sonar como el murciélago no puede percibir lo que percibe un organismo equipado con un sistema visual, como el hombre; puede presumirse que el mundo externo es muy distinto para uno y otro. De manera similar, aun cuando en menor escala, es presumible que la experiencia de dos individuos pertenecientes a la misma especie sea diferente. De acuerdo a esto, no parece errado decir -como abogados del argumento del vacío explicativo- que la ciencia cognitiva a menudo se presenta como una teoría de la mente que deja afuera

---

[1] Partita entre fenomenal, fenomenológico y fenomenología. La primera, utilizada como término general y amplio en las discusiones sobre la tercera y la primera persona, se refiere a cualquier cosa posible de ser examinada. Está definida en esta sección. "Fenomenológico" ha sido muy utilizado, en sentido general, para designar la experiencia consciente y la subjetividad (p.ej. por J. Jackendoff en sus trabajos pioneros y por alguno de los colaboradores en este número). Este uso amplio del término está ligado al anterior y debe ser claramente distinguido de trabajos basados en la tradición filosófica de la fenomenología, donde la descripción "fenomenológica" se basa necesariamente en la "reducción".

la fenomenalidad o la subjetividad, ya sea porque no intenta explicarla o porque no es capaz de hacerlo adecuadamente.

Estas ideas necesitan refinarse, sin duda, para llegar realmente al punto en cuestión. Y es aquí donde es necesario introducir la noción de conciencia. El progreso de la ciencia cognitiva (así como el desarrollo del psicoanálisis) nos ha familiarizado con la idea de que algo puede sucederle a un sujeto, y ser en este sentido subjetivo, pero quedar fuera del alcance del mismo sujeto. Normalmente describimos este caso diciendo que el sujeto no es consciente del fenómeno mencionado. Por lo tanto es necesario establecer una distinción entre fenómenos conscientes e inconscientes, o nuevamente, entre una subjetividad consciente y una sub-personal. La noción de conciencia o de estar consciente, en sí misma, designa en primer lugar el hecho —en relación con el fenómeno— de que el sujeto sabe de qué se trata, está informado.

Parece tentador combinar los conceptos de dato fenoménico y de conciencia subjetiva. Pero la noción de fenómeno no-consciente o subpersonal se opone a tal movimiento. Por ejemplo, hay momentos en que percibimos fenómenos de manera prerreflexiva, sin alcanzar a darnos cuenta de ello, pero en que un "gesto" o método de examen nos lo aclara o lo vuelve consciente. Lo que se cuestiona en este caso es la ingenua suposición de que la línea demarcatoria entre lo estrictamente subpersonal y lo consciente es inamovible.

No escapará a la atención del lector que todo este estudio en primera persona está basado en la conclusión (suposición) de que la experiencia vivida es irreducible, es decir, que los datos fenoménicos no pueden ser reducidos o derivados de la perspectiva de la tercera persona. Sin embargo, nos parece justo entregar al lector las tres bases principales que sustentan la suposición de irreductibilidad de la experiencia.

Primero, aceptar la experiencia como un campo por explorar es aceptar la evidencia de que la vida y la mente incluyen la dimensión de la primera persona, que es el sello de nuestra existencia tal como la vivimos. Dejar nuestra investigación científica fuera de

este campo fenoménico implica separar la vida humana de su dominio más íntimo, o negarle el acceso de la explicación científica. En ambos casos, el resultado será insatisfactorio.

Segundo, la experiencia subjetiva dice relación con el nivel del us*uario* de los propios saberes, o intenciones y acciones en las prácticas cotidianas. Sé que mis movimientos son el producto de una serie de contracciones musculares coordinadas. No obstante, la acción de mover mi mano opera en la escala emergente de los planos motores, que aparece ante mí como *intenciones* motoras, como un usuario-agente, no como tonos musculares que pudieran ser vistos de la posición de tercera persona. Esta dimensión práctica es la que permite la interacción con relatos de tercera persona en primer lugar (y no la descripción abstracta de un brazo de sofá tan familiar en la filosofía de la mente).

Tercero, en las prácticas humanas, la experiencia es el punto de partida privilegiado para producir cam*bios* con la mediación de profesionales de todo tipo, por ejemplo en la educación y el aprendizaje, el entrenamiento deportivo, la psicoterapia. En todos estos campos existe, por una parte, abundante evidencia de que la experiencia es esencial para la actividad humana y la vida incluyendo el uso de la propia mente, y por otra, de que el campo de la experiencia puede ser explorado, tal como vemos en las transformaciones conseguidas mediante prácticas específicas e interacciones humanas en situaciones prescritas (cursos de entrenamiento, preparación deportiva, sesiones de psicoterapia). Una vez más, es necesario poner en tela de juicio la supuesta fijeza de la línea de demarcación entre lo estrictamente subpersonal y lo consciente. Las metodologías de primera persona incluyen como una de sus dimensiones fundamentales la afirmación de que se trata de una línea móvil y que se puede hacer mucho en la zona de intermediación. La exploración de lo prerreflexivo representa una fuente muy poco explorada de información y datos, con enormes consecuencias.

4. La noción de método

A partir de aquí podemos comenzar esbozar lo que podría ser un método para todas estas tradiciones. En una primera aproximación podemos decir que hay al menos dos dimensiones que deben ser consideradas para que un método sea tal:

1) Que entregue un claro *procedimiento* para acceder a un campo fenoménico.

2) Que entregue medios claros de *expresión* y *validación* al interior de una comunidad de observadores que tengan familiaridad con los procedimientos señalados en 1.

Considerando que la distinción entre tener una experiencia (siguiendo un procedimiento) y la validación (siguiendo un intercambio intersubjetivo regulado) no es absoluta, la situación puede ser expuesta de la siguiente manera:

Método	Procedimiento	Validación
1 Introspección	Atención durante una tarea definida	Relatos verbales, mediatizados
2 Fenomenología	Reducción-suspensión	Variables descriptivas
3 Meditación: Samatha, Mahamudra, Zen, MT	Atención sostenida, conciencia expandida, suspensión de la actividad mental	Relatos tradicionales, informes científicos

Deberíamos por supuesto agregar que esta tabla es solamente una muestra del conocimiento actual sobre la materia, y que no se debería vacilar en la evaluación de las relativas fortalezas y debilidades de cada método. Por ahora y de manera tentativa podríamos hacerlo en los siguientes términos:

Evaluación de procedimiento	Evaluación de validación
1. Mediano, requiere perfeccionamiento	Mediano, buen uso de los protocolos
2. Mediano, requiere perfeccionamiento	Mediano, algunos ejemplos útiles
3. Bueno, métodos detallados	Históricamente rico, a veces se funda excesivamente en relatos internos, algunos buenos protocolos científicos, requiere más desarrollo

En resumen, nuestra conclusión general es que tenemos a la mano suficientes resultados para asegurar que estos métodos de primero persona no son una quimera.

5. Esbozo de una estructura común[2]

Contenido y acto mental

En lo ya expuesto de nuestra presentación, está implícito que todos estos métodos pueden ser legítimamente comparados, por cuanto poseen algunas estructuras nucleares comunes. Ahora intentaremos probar y entregar algunos indicadores de tales aspectos comunes.

Consideremos la situación de un sujeto en una situación experimental, al que se solicita que cumpla con una tarea precisa. Durante la realización de la tarea, la persona experimenta algo. Siguiendo a Vermersch, llamaremos a esto su contenido vivido, C1, como referencia para lo que sigue. En el contexto de la tarea, se le pide al sujeto (con o sin mediación) que examine (describa, analice, y/o se dé cuenta, atienda) a C1. Para que esto suceda, durante un corto tiempo la C1 será parte del contenido de una nueva experiencia: el examen de la propia mentación, un nuevo contenido, C2, producto

[2] Esta sección debe mucho al libro por aparecer que escribió Francisco Varela en colaboración con Natalie Depraz y pierre Vermersch. Nuestro agradecimientos a ambos por permitirnos presentar algunas de sus ideas en esta introducción (Depraz *et.al.*, por aparecer)

del acto de notificación de la propia mentación. C2 contará con un nuevo contenido adicional, caracterizado por la manera de acceso a C1. Nótese que para la aparición del par C1-C2 es necesaria una reorientación del pensamiento, una interrupción de la actitud natural, que normalmente no se detiene para acceder a sus propios contenidos[3].

El elemento central que aparece en esta estratificación, común a todos los métodos de primera persona, es la clara distinción entre *el contenido* de un acto mental (por ejemplo se me pide imaginar mi casa) y *el proceso* mediante el cual aparece tal contenido (cómo consigo producirla imagen que se me ha pedido). Desatender esta distinción fundamental es causa de mucha confusión. Para estar seguros, digamos que los métodos enunciados más arriba muestran algunas diferencias. Los primeros dos tipos implican una atención focalizada en el contenido, orientada a conseguir cierto conocimiento o visión. Por tanto, aquí, lograr el aprendizaje de un método es concomitante con la fácil internalización del desdoblamiento de C1 a C2. Los métodos del tercer y cuarto tipo son más sutiles. En las primeras etapas de entrenamiento de "meditación sentada" (*samatha*), el papel que desempeña la atención a los contenidos mentales inmediatos es crucial. Pero a medida que se desarrolla la práctica, uno puede desatender ese esfuerzo inicial y aquellos contenidos pasan a ser menos importantes que la calidad de la presencia consciente en sí misma, tal como indican las técnicas de *mahamudra-dzogchen*. La tradición védica de MT, *dzogchen* y budismo zen, y algunas prácticas de budismo *vajrayana*, son desde el inicio indiferentes a los contenidos (por encima y más allá de lo necesario para llevar a cabo el método mismo). Aquí el propósito es desarrollar el método hasta que se autoaniquile, produciendo estados de conciencia "pura" (sin contenidos).

Es preciso señalar que los métodos representados aquí, por lo

[3] Menos común aún es la posibilidad de pasar a una tercera etapa, donde C2 se convierte en el objeto de una experiencia, C2, la introspección de la introspección. Al contrario de lo que parece a primera vista, esto no indica una regresión infinita: dado que todo esto sucede en un breve lapso de tiempo, los niveles de examen se detienen allí, dado que la experiencia muestra claramente que ya la C2 implica un esfuerzo, y que para la C3 uno parece necesitar una atención sostenida altamente entrenada.

general, producen relatos en primera persona que parecen planos y pobres. Esta es una de las quejas más comunes de quienes critican los relatos en primera persona, y no están equivocados. Lo que falta es la continuación del proceso, que incluye una transformación de la actitud natural, a la que C2 apunta directamente. C2 pone en marcha una segunda fase que deja claro que desde la suspensión inicial, un nuevo campo vacío comienza a ser progresivamente llenado con datos fenoménicos. Se trata de la etapa de descubrimiento, de acumulación de experiencia, y esto, para cumplirse, requiere de una sostenida disciplina. Lo que desorienta es la aparente facilidad para acceder a la experiencia personal: la acumulación que la suspensión hace posible tiene su propio tiempo de desarrollo, que necesita un cultivo particular y una gran paciencia. En realidad, todos los métodos (y los practicantes individuales) tienen su propio tiempo de desarrollo, que puede variar ampliamente.

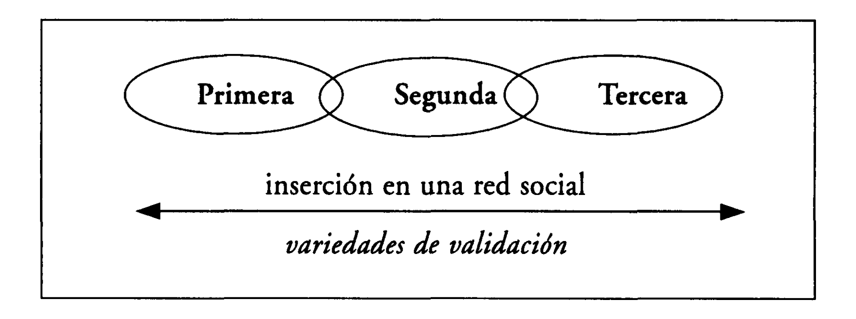
Segunda persona

Establecer un método requiere, por tanto, la creación de medios para sortear estas dificultades. Una vez más, encontramos que cada tradición ha creado sus propios medios y que éstos se diferencian claramente entre sí. No obstante, todos comparten un descubrimiento, cual es que el para familiarizarse progresivamente con un método es necesaria la mediación. Por mediación entendemos aquí otra(s) persona(s) que ocupa(n) una curiosa posición intermedia entre la primera y la segunda posición, de allí el nombre *segunda persona* para dicha posición. Un mediador es ex-céntrico respecto de la experiencia vivida C1 pero, no obstante, toma la posición de alguien que ha estado ahí hasta cierto punto y, por tanto, puede entregar consejos y entrenamiento.

La posición de segunda persona tiene gran importancia en todas las tradiciones, con la excepción de las ciencias naturales donde sólo aparece cuando se analiza el proceso social de aprendizaje y un investigador busca la mediación de un tutor más experimentado para fortalecer sus capacidades como científico. Aun cuando es común otorgar importancia al linaje científico de un investigador,

la atención sobre esta mediación de segunda persona desaparece cuando su artículo se publica en alguna revista científica.

En resumen, entonces, las tres posiciones (primera, segunda, tercera) se estructuran no tanto en relación con sus contenidos, sino más bien por la manera *en que* aparecen —insertas en la red de intercambios sociales. Son las funciones que los agentes sociales asumen en cada caso lo que determina su pertenencia a una u otra posición, con gradaciones. No estamos interesados, por tanto, en una oposición dual entre lo privado y lo público o entre lo objetivo y lo subjetivo. Lo que sí nos interesa son las cuestiones de interpretación de los resultados de las investigaciones de primera-persona, como datos válidos, y aquí el asunto de las relaciones entre las tres posiciones es de enorme importancia. Figurativamente, la situación puede ser expuesta de la siguiente manera:



Expresión y validación

Es precisamente porque la validación depende de la manera en que se constituye la red intersubjetiva, que estas posiciones, necesariamente, están marcadas por gradaciones. O, en otras palabras, la distinción de posiciones en tres fases es hasta cierto punto una convención descriptiva. Parece, en efecto, que cada una, en cada ocasión, aparece como una función del *énfasis* que uno pone en alcanzar un modo particular de validación. Reexaminemos las tres posiciones, de acuerdo a este nuevo enfoque.

La tercera persona, entendida como el observador estándar del discurso científico, enfatiza hasta el extremo el aparente dualismo entre lo interno y lo objetivo externo, y es la base del reduccionismo científico en todas sus formas. Las ciencias cognitivas representan un desafío intrínseco a esta forma "pura" de ciencia objetiva, dado que el tema de estudio implica directamente a los agentes sociales en función. Se trata, por tanto, de un caso singular, con características inencontrables en cualquier otro campo de las ciencias naturales. No obstante, considerando que sus contenidos científicos están focalizados en operaciones bioquímicas y neuronales, el problema de las personas (primera, segunda, tercera) no parece muy relevante. Esto se debe a que en los estudios sobre humanos, como en aquellos sobre otros animales, el comportamiento (incluido el comportamiento verbal) es una parte integral de los datos relevantes que pueden ser estudiados en el marco de los registros y mediciones clásicos, sin necesidad de ser considerados expresiones de la vida mental.

La imputación de un correlato mental de estos fenómenos verbales o de comportamiento es otro tema, por cierto, y el conductismo fundamenta su programa de investigación justamente en esta distinción entre conducta y subjetividad. En la práctica, la mayoría de los neurocientíficos hacen de esta distinción una cuestión de principios y asumen, de manera menos o más explícita, que tal conducta proviene de un agente o sujeto cognitivo, sin una elaboración más a fondo. Por lo tanto, aun cuando estas prácticas pertenecen plenamente a la norma de los relatos de tercera persona, ellas mismas insinúan las otras posiciones y temas, de manera muy distinta a como sucede en el estudio de las células y los cristales. Es por esto que distinguimos la ciencia cognitiva como aquella que implica, al interior de la tercera persona, una posición que colinda con la posición de segunda persona.

Desde una posición de segunda persona, lo que simplemente aparece como un comportamiento externo, es considerado explícitamente como huellas o manifestación de vida mental y, más aún, como la única vía que conduce a lo que es la vida mental misma. Recientemente, Daniel Dennet (1991) ha formulado esta posición llamándola con gran acierto h*etero*fenomenología, que

él mismo ilustra como la posición de un antropólogo estudiando una cultura remota. Para un antropólogo, la mera colección de grabaciones, fotos e inscripciones no dan forma a una verdadera respuesta. Lo que debe hacer es deducir modelos de vida mental (cultural) de estos materiales, trabajando desde una posición intencional. En otras palabras:

> *Debemos entender al emisor-de-ruidos como un agente, en realidad un agente racional que posee creencias y deseos y otros estados mentales que exhiben intencionalidad o definiciones previas acerca de cómo puedenb explicarse las acciones sobre la base de esos estados mentales (Dennet 1991, p.76)*

Pero el heterofenomenólogo no se abandona a la pura apariencia de su fuente de información, ni se queda con las interpretaciones que esta misma entrega. El antropólogo no se convierte en miembro de la tribu. La razón por la cual esta postura puede entenderse como una posición de segunda persona es que el heterofenomenólogo, aun cuando utilice señales estrictamente externas, está presente como un individuo situado, que debe generar interpretaciones intencionales de los datos. Nada de esto es necesario cuando la validación no considera ninguna forma de postura intencional, como hace el científico neuronal y conductista cuando estudia animales, y puede hacerlo cuando estudia casos humanos.

La tentación de convertirse en miembro de la tribu es precisamente donde nos lleva el próximo cambio de énfasis, esto es, la posición de segunda persona en sentido estricto. Aquí, en la posición de segunda persona, uno/una abandona de manera explícita su distanciamiento para identificarse con el tipo de entendimiento y coherencia interna de su fuente de información. De hecho, así es como entiende su papel: como un resonador empático con experiencias que le son familiares, que encuentran en él cierta resonancia. Esta posición empática es todavía en parte heterofenomenológica, dado que se necesita un mínimo de distanciamiento crítico y de evaluación crítica, pero la intención es enteramente distinta: producir un encuentro en el mismo terreno, como miembro del mismo tipo. En las tradiciones relacionadas con las prácticas humanas que hemos examinado se pueden encontrar abundantes

ejemplos de esta posición. No se trata de la posición neutra del antropólogo, sino más bien de aquella propia de un entrenador o una partera. Su tarea está fundada en una especial sensibilidad a las señales (siempre sutiles) del fraseo de su interlocutor, su lenguaje corporal y expresividad, en busca de los indicadores (menos o más explícitos) que incursionan en el campo de la experiencia común, tal como detallaremos más abajo. Tales encuentros no serían posible si el mediador no se adentrara en el campo de las experiencias por investigar, tal como nada puede reemplazar el conocimiento de primera mano. Este, por lo tanto, es un estilo de validación radicalmente distinto a los otros que hemos discutido hasta ahora.

Esta postura de segunda persona empática puede ser igualmente asumida desde el punto de vista de quien vive la experiencia, si fuera necesario. En este caso, la posición le compete al sujeto mismo, que progresa a través de la sesión del trabajo, con la posibilidad de buscar una validación a través de sus acciones y expresiones. Esta posible validación subjetiva no es siempre obligatoria. Pero, planteado a la inversa, nos encontramos con que no hay ninguna posibilidad de una metodología de primera persona, en el sentido que estamos entendiendo el término, sin asumir en algún momento la posición de la experiencia directa que busca validación. De otra manera el proceso se convierte en algo puramente privado e incluso solipsista. No negamos que una alternativa como esta es posibles para los seres humanos —incluso al extremo de mundos completamente autoimaginados, que desde el punto de vista panorámico de la red social aparecen como ilusiones y son tratados como tales. Pero dejando todas estas cuestiones de lado, el paso hacia una posición abierta a la mediación de segunda persona es claramente necesario para entender la envergadura de nuestra búsqueda metodológica.

Rasgos comunes

En resumen, hemos visto que los métodos de primera persona, como los hemos presentado aquí, comparten algunos rasgos o etapas comunes:

- *Actitud básica:* requieren de un momento de suspensión y cambio de dirección desde los contenidos a los procesos mentales.
- *Acumulación fenomenológica:* requieren de un entrenamiento específico para llevar la suspensión inicial a un contenido más lleno y el papel mediador de la segunda persona es aquí importante.
- Expr*esión y validación intersubjetiva*: En todos los casos, los procesos de expresión y validación requerirán de relatos explícitos sujetos a retroalimentación subjetiva.

También podemos concluir que las metodologías de primera persona no son cosa rápida ni fácil. Requieren de una dedicación sostenida y de un marco interactivo, antes que los datos fenomenológicos significativos puedan hacerse accesibles y validables. Finalmente, el papel de la mediación (que hasta ahora ha recibido escasa atención), es un aspecto único de estos métodos.

6. Grados de ceguera

La discusión sostenida más arriba nos permite ahora evaluar de manera más precisa la situación de las metodologías de primera persona en el contexto general de la ciencia cognitiva. En el actual debate sobre una ciencia de la conciencia hay una serie de autores que excluyen de manera explícita cualquier relación con los datos de primera persona: por razones de enfoque, estos autores no serán aquí considerados. Lo que nos parece interesante es echar una mirada a aquellos autores que sí se interesan por estos datos, y caracterizarlos por lo que podemos llamar su grado de ceguera ante el papel y la importancia de trabajar con métodos de primera persona.

Nuestra opinión es que el campo de los estudios sobre el conocimiento y la neurociencia cognitiva ha estado en demasía bajo la influencia de un estilo particular de filosofía de la mente, y completamente ajeno a otras tradiciones que se han especializado en la exploración metódica de la experiencia humana. Por lo mismo

no es sorprendente (salvo notables excepciones) que terminemos en la repetición recurrente de un mismo tipo de argumentos. Probablemente no hay manera más clara de ilustrar lo dicho que leyendo la guerra verbal entre David Chalmers y John Searle, recientemente recogida en el *New York Review of Books* (Searle, 1997). Lo que tenemos aquí es a dos filósofos que sostienen, cada uno a su manera, la necesidad de incorporar la dimensión subjetiva o experiencial como irreducible. Pero este reconocimiento de la importancia de la experiencia en primera persona no está acompañada por un progreso metodológico. No es sorprendente, en consecuencia, que Chalmers y Searle se acusen mutuamente de estar casi o completamente equivocados en todos los temas cruciales. Esto es una clara señal de la necesidad de buscar otros caminos de desarrollo, con nuevas fuentes y nuevas herramientas.

Veamos ahora la noción de "grados de ceguera" a la utilidad de trabajar con métodos de primera persona. El grado más elemental, por cierto, está representado por aquellos que son sensibles a la experiencia de primera persona pero desean acercarse a ella sólo por medio de la tercera persona. Esta posición es compartida por una gran mayoría de neurocientistas cognitivos y está bien articulada por Dennett en su noción de heterofenomenología.

Una segunda ceguera, más refinada, está representada por aquellos que sostienen que se debe considerar los relatos de primera persona, pero se detienen en esa afirmación y no proponen nada para llevar a cabo tal propósito. Esto se puede encontrar regularmente en la filosofía de la mente anglosajona, donde el problema del conocimiento se asimila a menudo al de "*qualia*" en relación con ciertas características o estados mentales. Tomemos por caso a Searle (1992), quien argumenta a favor de la irreductibilidad de la ontología de primera persona. Aun cuando llega al estudio de la subjetividad, no propone algo más que un rápido abandono de la introspección. Quiere que aceptemos que "la irreductibilidad de la conciencia es simplemente consecuencia de la pragmática de nuestras prácticas referidas a las definiciones" (p.122) y, por tanto, aun cuando la irreductibilidad del conocimiento es un "argumento claro, no tiene consecuencias profundas" (p.118):

El hecho mismo de la subjetividad, que estábamos tratando de observar, hace imposible tal observación. ¿Por qué? Porque allí donde está presente la conciencia subjetiva, no hay distinción entre el observador y la cosa observada... Cualquier introspección que tenga de mi propio estado de conciencia es ese mismo estado de conciencia (p.97).

La mente no tiene por tanto ninguna manera fiable de estudiarse a sí misma, lo que nos deja ante una conclusión lógica, pero en un limbo pragmático y metodológico[4]. Este limbo no se diferencia mucho del que ofrece Ray Jackendoff, quien a su manera también proclama la irreductibilidad de la conciencia pero, en cuanto a cuestiones de método, calla ostensiblemente. Sí sostiene que los enfoques sobre la experiencia constituyen contrapesos a una teoría computacional de la mente, pero no aporta ninguna recomendación metodológica, excepto "la esperanza de que el desacuerdo sobre la fenomenología pueda ser asumido en una atmósfera de mutua confianza" (Jackendoff, 1987, p. 275).

En un tercer nivel de inclusión, algunos han sido algo más constructivos. Chalmers, por ejemplo, sostiene que "un enfoque [fenomenológico] debe ser absolutamente central en una adecuada ciencia del conocimiento" (Chalmers, 1997, p.36), y considera que tanto las metodologías orientales como las occidentales pueden llegar a ser útiles. Pero su propio trabajo, hasta ahora, no parece haber ido más allá, en cuanto metodologías, de esta positiva sugerencia. La noción de equilibrio reflexivo de Owen Flanagan combina los relatos fenomenológicos y de tercera persona en una suerte de "método

---

4 No obstante, es un hecho que los estados mentales pueden ser examinados y que las teorías acerca de ellos pueden ser aceptadas y/o rechazadas. En rigor, los estudios de desarrollo cognitivo muestran que los niños deben aprender cuáles de sus experiencias pueden ser situadas en el mundo objetivo y cuales sólo tienen lugar en sus mentes (incluso los adultos pueden llegar a confundir estos planos). Esto hace posible pensar que hay una forma de "observación" que es anterior e independiente de la bifurcación subjetivo/objetivo que, según Searle, hace imposible la observación subjetiva (ver Shear, 1996).

natural" que conjuga datos fenomenológicos, psicológicos y neuronales (Flanagan, 1994, p.11). Podemos coincidir con él y seguir algunos de sus ejemplos (sobre audición dividida, o percepción bi-estable). Pero este análisis de caso-por-caso dista de una metodología sistemática, que Flanagan no desarrolla. Bernard Baars es notablemente más explícito al introducir un método que llama "fenomenología contrastiva". Señala: "La clave es comparar dos procesos cerebrales activos similares en muchos aspectos, pero diferentes en relación con la conciencia" (Baars, 1997a, p.21). Este es un procedimiento básico, implícitamente considerado en la mayoría de los estudios empíricos que hoy se publican sobre correlatos neuronales de los procesos conscientes, tales como la atención y la imaginación. Aun cuando sea este un importante paso adelante en metodología, el trabajo con las condiciones contrastivas deja todavía una gran cantidad de datos fenoménicos subjetivos por explorar. Los métodos discutidos hacen evidente que se puede explorar mucho más a fondo en la propia experiencia, si se hace uso de un método capaz de suspender de manera consistente los juicios y los pensamientos habituales[5].

Para concluir, digamos que este rápido repaso de los grados de ceguera muestra que hay todavía un largo camino para llegar a incorporar plenamente las metodologías de primera persona en la ciencia cognitiva. Hay, no obstante, señales de que esto lentamente comienza a suceder[6].

---

[5] Incluso para un escritor como Baars, esta necesidad de un desarrollo posterior provoca resistencias. En la discusión sobre un resumen de su libro, uno de nosotros puso el acento exactamente sobre este punto (Varela, 1997), lo que no fue bien recibido por el autor. Dice Baars en su réplica: "Todo el revuelo acerca de los métodos es todavía reducible a '¿es usted consciente ahora mismo de esta palabra que tiene ante los ojos?'. No es necesario complicar tanto las cosas" (Baars, 1997b, p.375). No obstante, hay una gran dosis de complejidad que no queda revelada en los datos de primera persona, y es aquí donde aparece la necesidad de métodos más refinados.

[6] Se debe reconocer aquí el magnífico trabajo pionero de Eugene Gendlin, que merece conocerse mejor (Gendlin, 1972/ 1997). En Velmans (en prensa) se puede encontrar una reciente y fructífera recolección de metodologías comparativas.

7. ¿PUEDE EXPLORARSE LA EXPERIENCIA? UNA OBJECIÓN INICIAL

Puede ser que nuestra orientación pragmática deje fríos a algunos lectores. El lector más escéptico deseará poner en el tapete lo que quizás es la objeción fundamental a lo que hemos venido proponiendo, y puede hacerlo de la siguiente manera: ¿Cómo puede uno saber si, al explorar la experiencia con un método, no está, de hecho, deformando o incluso creando lo que experimenta? Siendo la experiencia lo que es, ¿cuál es el posible significado del examen? Podemos llamar a esto una "falacia de excavación" o, filosóficamente hablando, la objeción hermenéutic*a* que apunta al corazón de nuestro proyecto. También se la puede encontrar como la objeción *deconstructiva*, basada en los análisis filosóficos posmodernos (casi todos derivados de J. Derrida). Aquí el énfasis recae en el argumento de que no hay algo como un plano "más profundo" de experiencia, dado que cualquier relato está "siempre ya" (*toujours déjà*) envuelto en el lenguaje, por lo cual un nuevo relato sólo puede ser una inflexión de prácticas lingüísticas.

Nuestra respuesta a la falacia de la excavación/deconstrucción es, en primer lugar, admitir que hay aquí un problema importante y que ninguna contorsión metodológica o argumento *a priori* lo resolverá *per se*. Parece imposible evitar que el método sea siempre parte de los tipos de entidades y propiedades que se encuentran en el campo de observación considerado. Tal como quedó dicho en la sección I (más arriba), la dimensión experiencial y social de la ciencia está a menudo escondida, pero nunca enteramente ausente. Esto es evidente en las formas más consagradas de la ciencia natural, y así lo han dejado en claro estudios recientes (ver los últimos estudios de Shapin y Shaeffer [1994] sobre la bomba de aire de Boyle).

En efecto, ninguna aproximación metodológica a la experiencia es neutra: inevitablemente se introduce un marco interpretativo en la reunión de datos fenoménicos. En la medida en que esto sucede, la dimensión hermenéutica del proceso es inevitable: todo examen es una interpretación, y toda interpretación revela y oculta al mismo tiempo. Pero esto no significa que una disciplinada aproximación a

la experiencia sólo entregue resultados artificiales o una versión "deformada" de lo que "realmente" es la experiencia.

Para ser claros, digamos que la exploración de la experiencia sufrirá las expectativas culturales y los sesgos instrumentales como cualquier otra investigación metodológica, pero no hay evidencia de que los datos fenoménicos reunidos no estén igualmente limitados por la realidad propia de los contenidos de conciencia. Por lo tanto, todas las descripciones que podamos producir mediante métodos de primera persona no serán descripciones puras, ni "hechos" sólidos, sino más bien ítemes de conocimiento intersubjetivo potencialmente válidos, quasi-objetos de tipo mental. Ni más ni menos.

Más aún, la experiencia humana no es campo fijo, delineado de antemano. Por el contrario, es cambiante, variable y fluido. Si uno se entrega a un entrenamiento disciplinado en interpretación musical, las nuevas capacidades en la distinción de sonidos, en la sensibilidad al fraseo musical y la interpretación grupal, serán innegables. Pero esto significa que la experiencia ha sido explorada y modificada mediante procedimientos disciplinados y nunca arbitrarios. En rigor, no tiene sentido hablar de la experiencia como algo estándar, tosco o puro. Lo que tenemos es siempre experiencia a su propio nivel de examen y dependiente del tipo de esfuerzo y método puesto en juego. Se mueve, cambia y su exploración es ya parte de la vida humana, aun cuando, por lo general, con otros objetivos que la comprensión de la experiencia misma.

Queremos por lo tanto situarnos en un campo intermedio respecto de la objeción hermenéutica. Por una parte queremos explorar al máximo las herramientas con las que contamos para realizar informes de primera persona. Por otra, no sostenemos que este esfuerzo esté libre de limitaciones metódicas o sea natural en algún sentido que lo privilegie. Esta combinación es otra manifestación del espíritu pragmático del trabajo necesario en primera persona. El tiempo dirá si esta orientación lleva en sí los frutos que esperamos. No se gana nada con rechazar la empresa completa por algunos argumentos *a priori*.

Baars, B.J. (1997a), *In the Theater of Consciousness* (Nueva York: Oxford University Press).

Baars, B.J. (1997b), 'Reply to commentators' *Journal of Consciousness Studies*, 4 (4), pp.3 47-64.

Chalmers, D.J. (1997), 'Moving forward on the problem of consciousness', *JCS* 4 (1), pp. 3-46.

Dennett, D.C. (1991), *Consciousness Explained* (Nueva York: Little, Brown).

Depraz, N., Varela, F. y Vermersch, P., *On Becoming Aware: Steps to a Phenomenological Pragmatics*.

Flanagan, O. (1994), *Consciousness Reconsidered* (Cambridge, MA: MIT Press).

Gendlin, E. (1972/1997), *Experiencing and the Creation of Meaning* (Evanston, IL: Northwestern University Press).

Jackendoff, R. (1987), *Consciousness and the Computational Mind* (Cambridge, MA: MIT Press).

Roy, J.M., Petitot, J., Pachoud, B. y Varela, F. (1998), 'Beyond the Gap. An Introduction to naturalizing phenomenology', en *Naturalizing Phenomenology: Current Issues in Contemporary Phenomenology and Cognitive Sciencie*, ed. Petitot et. al. (Stanford University Press; en imprenta).

Searle, J.R. (1992), *The Rediscovery of the Mind* (Cambridge, MA: MIT Press).

Searle, J.R. (1997), *The Mistery of the Consciousness* (Nueva York: New York Review of Books).

Shapin, S. y Shaeffer, S. (1994), *Leviathan and the Air Pump* (Princeton, NJ: Princeton Univ. Press).

Shear, J. (1996), 'The hard problem: Closing the empirical gap', *JCS*, 3(1), pp. 54-68.

Shear, J. (ed. 1997), *Explaining Consciousness: The Hard Problem* (Cambridge, MA: MIT Press).

Varela, F. (1996), 'Neurophenomenology', *JCS*, 3 (4), pp. 330-49.

Varela, F. (1997), 'Metaphor to mechanism: natural to diciplined', *JCS*, 4(4), pp.344-6.

Velmans, M., *Investigating Phenomenal Consciousness: New Methodologies and Maps* (Amsterdam: Benjamins).

CONCIENCIA DEL TIEMPO PRESENTE

El tiempo es la fuente misma de la experiencia y la conciencia. Por lo mismo una exploración neurofenomenológica detallada de la temporalidad vivida y natural representa el test más exigente para mirar de cerca la fertilidad y viabilidad de mi posición sobre el tema de la conciencia. Por ende, es un texto arduo, pero donde el lector podrá encontrar una riqueza de detalle, que discusiones más generales no permiten.

1. Introducción

Mi objetivo en este artículo[1] es proponer una explicación explícitamente naturalizada de la experiencia del ahora presente sobre la base de dos fuentes complementarias: el análisis fenomenológico y la neurociencia cognitiva. A medida que el artículo avance se aclarará lo que quiero decir con el término naturalización, así como el papel que cumple la neurociencia cognitiva, pero mi principal intención es utilizar la conciencia del tiempo presente como un *estudio de caso* para análisis neuro-fisiológico.

Sería una insensatez afirmar que se puede abordar este tema y agotarlo. La experiencia de la temporalidad plantea frontalmente el

---

(Varela, F.J. (1999), Present-time consciousness, *J.Consc. Studies Studies* 6 (2-3):111-140.)

1 Este artículo es una adaptación de "The specious present: The neurophenomenology of inner time consciousness" en *Naturalizing Phenomenology*, editada por Petitot, Varela, Pachoud y Roy, que será publicado próximamente por la Stanford University Press. Agradezco a la Stanford University Press por autorizar esta versión.

hecho fundamental de que existimos sólo dentro de una red transparente de tiempo. En general, su elucidación ocupa un lugar central en la historia del pensamiento y, más certeramente, en la tradición fenomenológica. Edmund Husserl consideraba la temporalidad como un eje fundacional de su investigación fenomenológica: todas las demás formas de actividad mental dependen de la temporalidad, pero ésta no depende de ellas. El estudió estos problemas hasta su muerte. A diferencia de sus muchos predecesores ilustres (incluido James), Husserl logró provocar un progreso esencial en la formulación de las estructuras básicas del tiempo íntimo (Gallagher, 1998). Inspirarse en la explicación fenomenológica desarrollada por Husserl no representa ningún tipo de obsesión husserliana escolástica. No está de más insistir en que al usar la utilización de Husserl de la fenomenología del tiempo no estoy preocupado de realizar una detenida lectura textual con el fin de aprobar o desaprobar una parte del pensamiento del autor. Prefiero seguir el ejemplo del *estilo* de Husserl como un eterno principiante, siempre dispuesto a comenzar de nuevo; este es el sello de la fenomenología misma (pero, en la práctica, éste no siempre ha sido el caso). Lo que más me interesa es el movimiento inconcluso de sus escritos: sólo en sus giros y saltos conseguiremos llegar a ver cómo su descripción contiene, entre muchas otras ideas en germen, una aptitud dinámica que es clave para un proyecto de naturalización.

Estos datos fenoménicos facilitarán los puentes para la neurociencia cognitiva, tratada aquí en relación con los recientes resultados que provienen de la dinámica neural, y la emergencia de unidades a gran escala, que dan un contrapunto a la constitución de la temporalidad. Mi enfoque de la temporalidad es un estudio de caso de una dirección de investigación general que he llamado *neurofenomenología*, en la cual la experiencia vivida y su base biológica natural están vinculadas a través de *restricciones mutuas* proporcionadas por sus descripciones respectivas (Varela 1996). Dada la importancia del tema de la experiencia de la temporalidad, quiero dejar claro que considero esto como una prueba mordaz de toda la empresa neuro-fenomenológica.

2. El tiempo vivido no es físico-computacional

Como es común en cualquier estudio fenomenológico verdadero, el estudio del tiempo implica el gesto de la reducción[2], seguido por la identificación de invariantes descriptivas. Tan pronto como entramos en un estudio del tiempo presente, con este tipo de actitud filosófica, se hace evidente que la conocida explicación del tiempo presente heredada de nuestro moderno trasfondo cultural occidental es inadecuada. De hecho, hemos heredado de la física clásica una noción de tiempo semejante a una flecha de momentos infinitesimales, que fluye en una corriente constante. Se basa en secuencias de elementos finitos o infinitesimales, que incluso son reversibles para gran parte de la física. Esta visión del tiempo es completamente homóloga a la desarrollada por la teoría moderna de la computación. Como una expresión refinada de la computación general, una máquina de Turing y su cabezal de escritura inscribe símbolos uno a uno en una sucesión infinita, dando origen al tiempo como una secuencia, exactamente como en la mecánica clásica.

A medida que las visiones computacionales entraban en la ciencia cognitiva a través del punto de vista computacionalista (o cognitivista), el tiempo computacional fue utilizado incondicionalmente en el estudio cognitivo del tiempo. Algunas investigaciones continúan basando los estudios experimentales en un "reloj interno" que da origen a una duración en varias escalas. Un reloj (hipotético) emite pulsaciones que se traducen en conducta; los juicios acerca de la duración dependen del conteo de las pulsaciones y se reflejan en la memoria y la decisión (Church y Broadbent, 1990). Esta estricta adherencia al esquema computacional será, de hecho, uno de los marcos de investigación que necesita ser aba*ndonado* como resultado del examen neuro-fenomenológico aquí propuesto. Pero volveré más tarde a esta conclusión, después de presentar mi argumento principal.

Incluso bajo una reducción superficial, ya proporcionada por reflexiones como las de Augustine y James, el tiempo en la *experiencia*

2 En este artículo, el término "reducción" y sus cognados son utilizados en su sentido técnico husserliano y no con su significado más común en la filosofía de la ciencia.

es algo completamente diferente a un reloj en tiempo lineal. En primer lugar, no se presenta a sí mismo como una secuencia lineal, sino como una *textura* compleja (de ahí que el "tiempo engañoso" de James[3], no es un presente "bien marcado"), y su plenitud es tan extraordinaria que domina nuestra existencia en un grado importante. En una primera aproximación, esta textura puede ser descrita de la siguiente manera: siempre existe un centro, el momento actual con un contenido intencional centrado (por ejemplo, esta habitación con mi computador frente a mí en el que las letras que estoy escribiendo son destacadas). Este centro está limitado por un horizonte o margen que ya es pasado (aún guardo en mi mente el comienzo de la oración que acabo de escribir) y se proyecta hacia un próximo momento buscado (esta sesión de escritura aún no termina). Estos horizontes son móviles: este mismo momento que era presente (y, por consiguiente, no fue meramente descrito, sino vivido como tal) se desliza hacia un presente pasado inmediato. Entonces, se sumerge profundamente y desaparece de nuestra vista: no lo guardo en mi mente de manera tan inmediata y necesito una profundidad adicional para mantenerlo a mano. Esta textura básica es la materia prima de lo que será tratado *in extenso* más adelante. En líneas generales, nos referiremos a esto como la ***estructura de tres partes de la temporalidad.*** Esto representa uno de los resultados más notables de la investigación de Husserl con la reducción fenomenológica.

Otro aspecto complementario importante de la temporalidad, tal como aparece en la reducción, es que la conciencia no contiene el tiempo como una categoría psicológica constituida. En cambio, la conciencia temporal *misma* constituye un sustrato primordial de la conciencia en que no se puede realizar una mayor reducción, un

> *"medio universal de acceso a lo que exista... La fenomenología constitutiva bien puede ser caracterizada como el desarrollo constante y radical de este privilegio de la conciencia en sus consecuencias y ramificaciones últimas" (Gurwitsch, 1966, p.xix).*

3 La expresión es de James. De hecho, Husserl recalcó cuidadosamente en su copia de *Principles* el término "presente engañoso" y sugirió, en taquigrafía de Gabel, una traducción en alemán ¡que permanece sin decifrar! Agradezco a Shaun Gallagher por informarme de esto.

Encontramos una conclusión convergente en James respecto a la paradoja aparente de la experiencia humana temporal: por una parte, existe la unidad del presente, un conglomerado que podemos describir como el lugar en que residimos en la conciencia básica y, por otra, este momento de conciencia es inseparable de un *flujo*, una corriente (Capítulo IX de *Principles*, 1898). Estos dos aspectos complementarios de la conciencia temporal son los ejes principales de mi presentación.

Este análisis preliminar aproximado de la conciencia del tiempo lleva, así, a tres niveles distinguibles de temporalidad que guiarán mi argumento:

(1) Un primer nivel, propio de los *objetos y sucesos temporales* en el mundo. Por lo tanto, se encuentra cerca de las nociones comunes de temporalidad en la experiencia humana, que fundamenta aquella actualmente utilizada en física o computación, pero también en la psicología experimental del tiempo.

(2) El fenomenólogo comienza por este nivel, pero la reducción hace que el segundo nivel sea evidente, el de los *actos de conciencia* que constituyen los objetos-sucesos[4]. Este es el tiempo "inmanente" o "tiempo interior" de los actos de conciencia. Su estructura da forma al cuerpo principal del análisis fenomenológico en las *Lecturas* de Husserl.

(3) Finalmente, (y este es el nivel de análisis más sutil), estos primeros dos niveles están constituidos por otro nivel en que no es posible ninguna distinción interna-externa, y que Husserl denomina el "*flujo de la conciencia* que constituye el tiempo absoluto" (PZB 73)[5].

4 En este artículo prefiero utilizar la expresión conjunta objeto-suceso para traducir *Zeitobjekte*. Aunque la terminología utilizada por G. Granel (1968), tempo-objeto sería mejor, es gramaticalmente incorrecta.

5 La sección o número de página corresponde a *Zur Phänomenologie des Inneren Zeitbewusstseins, Husserliana X* (1966), abreviada aquí PZB. La traducción al inglés "absolute time constituting flow of consciousness" proviene de los textos de J. Brough, J. Churchill y R. Bruzina cuando aparecía en ellos, o bien los he modificado o proporcionado la propia.

3. La duración de los objeto-sucesos

3.1. Duración: La experiencia de la multiestabilidad visual

El tiempo nunca aparece distanciado, sino como objeto-sucesos que son los correlatos o el foco intencional de la conciencia temporal: los objeto-sucesos temporales constituyen el *contenido* de estos actos. A diferencia de lo que podría ser interesante para el sicólogo o el científico neurólogo, para el fenomenólogo el contenido del objeto no es tan importante como su forma de aparición.

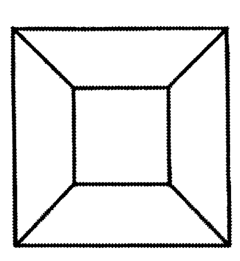
En varios puntos de su investigación Husserl vuelve a la observación básica de que lo que es propio de los objetos temporales es su doble aspecto de d*uración y unidad* (PZB 23, 113-14). La duración es correlativa con la dirección intencional: *esta* casa por la que camino, *ese* pájaro que voló de aquí hasta allá... Estas son duraciones reales, y se refieren al objeto que tiene una ubicación en el tiempo. La unidad es correlativa con la individualidad de los objeto-sucesos en cuestión, que se opone como un todo distinto al trasfondo de los otros sucesos. Así, un objeto-suceso temporal cubre un período T, un acto completo. Sin embargo, todo el acto es un proceso continuo en el curso del cual se *articulan momentos del ahora*, no como una unidad terminada, sino como en una sucesión. Lo que me interesa es este modo de constitución de los objeto-sucesos, cualquiera sea su duración T y su contenido. Debe ser así que la conciencia de sucesión deriva de las características estructurales de los actos de conciencia. Nuestro problema es la caracterización de estas estructuras.

Es característico que Husserl dé pocos ejemplos en sus obras. En PZB utiliza un ejemplo recurrente: escuchar una melodía (una elección que, más tarde, los comentaristas también han seguido). Por varias razones, creo que es importante estudiar este problema con más atención. En primer lugar, porque cuando Husserl da su ejemplo es, aunque parezca extraño, como un sujeto sin ubicar de manera abstracta: no sabemos en qué circunstancias la música está siendo escuchada (si él está solo o en una sala de conciertos), ni si es parte del contexto o está siendo escuchada con una concentración

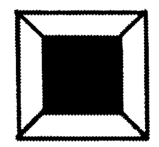
emocional intensa (¿es una pieza emotiva, es conocida por él?). Todo esto no es sólo periférico, puesto que sin estas particularidades se pierde el modo de acceso a la experiencia misma.

UNA TAREA DE PERCEPCIÓN VISUAL

- Observe la siguiente imagen:



Fuera de contexto, la mayoría de los sujetos occidentales naturalmente ven ya sea una "pirámide" o un "pasillo":





- Suspenda todas las interpretaciones habituales y considere las múltiples variaciones en conjunto. Podemos inferir que como no representa ni una pirámide ni un pasillo, es una percepción ambigua o multi-estable.
- Estas variaciones pueden ser prac*ticadas* mediante la siguiente estrategia:
  1. Fijar la atención en el centro de la figura.
  2. Parpadear.
  3. "Dirigirse" a su alternativa.

Aún extrañamos una fenomenología de conciencia-tiempo interno en que los gestos reduccionistas y la base textural de la experiencia figure explícita y totalmente. Toda la fenomenología de la conciencia-tiempo interno debería ser re-escrita con la precisión de un modo de acceso a la experiencia que sirva como base para la reducción; lo que sucede en la obra de Husserl está, a menudo,

alejado de esta textura. Aquí me aparto del uso, invitando al lector a entrar realmente en una experiencia específica de percepción visual multi-estable que dará al lector y al autor un *fundamento fenoménico común* explícito. Las percepciones multi-estables son un buen estudio de caso: son lo suficientemente precisas y complejas para el análisis, y, además, entregan los correlatos neurocognitivos correspondientes. Además, los fenómenos de percepción visual multiestable son aparentemente claros.

Estas instrucciones transforman al sujeto en un agente activo, como la condición de posibilidad de la percepción, la constitución del significado mismo ilustrada aquí en un caso mínimo. Hay mucho más que estudiar en este experimento en lo relativo al horizonte perceptual utilizado para cambiar la percepción, así como con respecto a la estrategia del lenguaje (el nombre de la alternativa, el "dirigirse" a las otras posibilidades)[6]. Pero permítanme seguir con los aspectos temporales de esta percepción. Nos damos cuenta, como se sabe desde los gestaltistas, de que, a medida que adquirimos competencia, el cambio puede ser llevado a cabo con poco esfuerzo y bajo solicitud. También está claro que el cambio tiene en sí mismo una dinámica muy compleja que se encarga de una "vida" propia (más adelante me referiré a esto). También advertimos que el gesto del cambio va acompañado de una "profundidad" en el tiempo, una duración *incomprimible* que hace que la transición se perciba como un cambio repentino de un aspecto a otro y no como una secuencia progresiva de cambios lineales.

Estamos de aceurdo que este conocido fenómeno no es común en la vida ordinaria. Un caso más "ecológico" sería, por ejemplo, cuando abro una puerta y al cruzar el umbral me encuentro con alguien que está directamente frente a mí. Tanto a través del movimiento del cuerpo, como por la orientación visual, el rostro de la persona se comienza a ver espontáneamente con claridad: reconozco a un colega y extiendo mi mano para saludarlo. Quiero destacar lo que sucede cuando uno se aleja de los objeto-sucesos dominado por

6 Ver por ejemplo Ihde, 1977, cap.6.

una actitud pasiva como al escuchar música, tanto en la percepción visual multiestable, como en la escena recién relatada.

En el ejemplo de las percepciones multiestables no necesitamos mover activamente todo nuestro cuerpo. Sin embargo, un importante estrato de movimiento está presente activamente, ya sea en el ajuste de la cabeza, el fruncir el ceño y el parpadeo y, por cierto, en los diferentes movimientos de los ojos. Esto es importante, puesto que no se puede concluir mucho de los casos en que el movimiento está ausente. Tal como la investigación fenomenológica misma ha repetidamente enfatizado, la percepción se basa en la interdependencia sensori-motriz activa[7]. Independientemente, varias tradiciones en la investigación cognitiva han subrayado, a su modo, la percepción-acción como una clave[8]. Es este lado *activo* de la percepción el que da a la temporalidad sus raíces en la vida misma. En este marco general, me concentraré con más precisión en la base estructural y las consecuencias de esta integración sensori-motriz para nuestra comprensión de la temporalidad.

3.2. La neurodinámica de la aparición temporal

Todo mi enfoque de la cognición se basa en agentes situados, encarnados. Para designar con más precisión este enfoque, he introducido el nombre de *enactivo*. Este se compone de dos aspectos complementarios.

(1) Por una parte, el acoplamiento en curso del agente cognitivo, un acoplamiento permanente que es logrado fundamentalmente a través de actividades sensori-motrices.

(2) Por otra, las actividades autónomas del agente, cuya identidad se basa en las configuraciones endógenas emergentes (o patrones de autoorganización) de la actividad neuronal.

---

7 Comúnmente esto se atribuye al trabajo de Merleau-Ponty, pero es muy explícitamente articulado por Husserl en materiales sólo recientemente disponibles. Ver Zahavi (1994).

8 Algunos clásicos son E. von Holst, J. Piaget y j.j. Gibson. Para un análisis reciente ver Varela *et al.* (1991); Kelso (1994).

La enacción implica que el acoplamiento sensori-motriz modula, pero no determina, una actividad endógena en curso que configura en elementos significativos del mundo en un flujo incesante.

No puedo ampliar mucho más este marco general[9], pero es el telón de fondo de mi análisis de la temporalidad como un proceso neurocognitivo. La enacción está naturalmente enmarcada en los instrumentos derivados de los sistemas dinámicos, en marcado contraste con la tradición cognitivista que encuentra su expresión natural en los modelos sintácticos de procesamiento de la información. El debate que enfrenta a la dinámica encarnada con los modelos computacionales abstractos como base de la ciencia cognitiva aún sigue vigente[10]. Por algún tiempo he abogado por el primero y en contra del segundo, y esta elección justifica la amplia utilización de los instrumentos dinámicos en este artículo.

Desde un punto de vista enactivo, cualquier acto mental se caracteriza por la participación simultánea de varias regiones del cerebro, funcionalmente distintas y topográficamente distribuidas, y su encarnación sensori-motriz. Desde el punto de vista del neurocientífico, la compleja tarea de relacionar e integrar estos diferentes componentes es el fundamento de la temporalidad. Una idea seguida aquí es que estos muchos componentes requieren un marco o v*entana de simultaneidad que corresponde a la duración del presente vivido.* En esta visión, el flujo constante de activación sensorial y resultado motriz es incorporado dentro del marco de una dinámica endógena (no una computacional-informática), que le da su profundidad o incomprensibilidad. Esta idea no es simplemente una abstracción teórica: es esencial para la comprensión de una vasta serie de evidencias y predicciones experimentales[11].

9 Ver Varela (1979); Varela *et al.* (1991); Thompson y Varela (1999).

10 Port y Van Gelder (1995), entregan un excelente estudio de los problemas contemporáneos, que incluye un asequible debate acerca de los instrumentos dinámicos para los científicos cognitivos.

11 Cf. Varela et al. (1981); Dennett y Kinsbourne (1991); Pöppel (1988); Pöppel y Schill (1995).

Estos marcos integrativos constituidos endógenamente explican el tiempo percibido como discretizado y no lineal, puesto que la naturaleza de esta discretización es más un horizonte de integración que una sucesión de '*quanta*' temporales.

En este punto, es importante introducir *tres escalas de duración* para comprender el horizonte temporal tal como acaba de ser presentado:

(1) sucesos básicos o elementales (la escala "1/10");

(2) período de relajación para integración a gran escala (la escala "1");

(3) afirmaciones descriptivo-narrativas (la escala "10").

Esta estructuración recurrente de las escalas temporales compone un todo unificado y sólo tiene sentido en relación con los objeto-sucesos. Plantea la pregunta de cómo algo temporalmente extendido puede aparecer como presente y también extenderse en mi horizonte temporal. La importancia de esta jerarquía recurrente de tres niveles aparecerá a lo largo de todo este artículo.

El primer nivel ya es evidente en el denominado intervalo de fusión de diversos sistemas sensoriales: la distancia mínima necesitada para que dos estímulos sean percibidos como no simultáneos, un umbral que varía con cada modalidad sensorial[12]. Estos sucesos elementales pueden fundarse en los ritmos celulares intrínsecos de las descargas neuronales, y en las capacidades de sumatoria temporal de la integración sináptica. Ellos entran en un rango de 10 milisegundos (por ejemplo, los ritmos de las descargas de interneuronas) a 100 mseg (por ejemplo, la duración de una secuencia EPSP/IPSP en una neurona cortical piramidal). Estos valores son la base de la escala 1/10. Conductualmente, estos sucesos elementales dan origen a fenómenos micro-cognitivos estudiados de forma muy diversa como momentos perceptuales, oscilaciones centrales, memoria icónica, ciclos de excitabilidad y cuantía de

12 En el Capítulo XV de *Principles*, James (1898) entrega una elegante descripción de estos datos que fueron ampliamente estudiados en el siglo XIX.

tiempo subjetivo. Por ejemplo, bajo condiciones estacionarias mínimas, el tiempo de reacción o conducta óculo-motriz muestra una distribución multimodal con una distancia de 30-40 mseg entre los puntos máximos; con una intensidad luminosa promedio, el movimiento evidente (o "psi-fenómeno") requiere de 100 mseg.

Esto nos lleva naturalmente a la segunda escala, la de la integración a largo plazo. Los procesos constituyentes tienen corta duración, del orden de los 30-100 mseg; ¿cómo se pueden comprender estos resultados experimentales sicológicos y neurobiológicos a nivel de una operación cognitiva normal y totalmente constituida? Una antigua tradición en las neurociencias estudia las bases cerebrales de los actos cognitivos (acción-percepción, memoria, motivación y otros) en términos de *asambleas celulares* o, análogamente, *conjuntos neuronales.* Una asamblea celular (AC) es un subconjunto de neuronas distribuido con fuertes conexiones recíprocas[13].

En el lenguaje de la dinámica, la AC debe tener un *periodo de relajación* seguido por una bifurcación o fase de transición, es decir, un período de emergencia dentro del cual surge, florece y se apaga, sólo para comenzar otro ciclo. Este período dilatorio está delimitado por dos limitaciones simultáneas (1) debe ser mayor que el periodo de los sucesos elementales (la escala 1/10); (2) debe ser comparable al tiempo que toma terminar un acto cognitivo, es decir, del orden de unos pocos segundos, (la escala 1), (por ejemplo, Varela *et al.*, 1981; Pöppel, 1988). En resumen, como dijimos antes, los procesos cerebrales pertinentes para la actividad cognitiva en curso no sólo están distribuidos en el espacio, sino también en una extensión de tiempo que no puede ser comprimida más allá de cierta fracción de segundo, la duración de la integración de los sucesos elementales.

¿Qué clase de evidencia hay para postular que todo acto cognitivo, desde la conducta percepto-motriz al razonamiento humano, surge de la actividad coherente de una subpoblación de

---

13 Esta sección proviene substancialmente de una publicación previa (Varela, 1995) que trata de manera más extensa las obras y la historia de estas ideas.

neuronas que se encuentran en múltiples ubicaciones? Y además, ¿cómo, estas asambleas, son momentáneamente autoseleccionadas para cada tarea específica? Como esto será un tema recurrente en lo que queda de este artículo, quisiera formularlo como una hipótesis de trabajo. La intuición básica que proviene de este problema es que una AC específica emerge a través de un tipo de *resonancia* temporal o "gluten". Más específicamente, el proceso de coherencia-generación neural puede ser comprendido de la siguiente manera: *Una AC específica es seleccionada a través de la efímera y rápida fase que bloquea las neuronas activadas que pertenecen a un sub-umbral, y que compiten por las AC.*

La idea clave aquí es que los conjuntos surgen debido a que la actividad neural forma conglomerados momentáneos de señales *de fase bloqueada* que provienen de múltiples regiones. *Forzosamente* debe producirse la sincronía (vía bloqueo de fase), a una velocidad lo suficientemente alta para que haya suficiente tiempo para que el conjunto "se mantenga" unido dentro de las limitaciones de los tiempos de transmisión y los marcos cognitivos de una fracción de segundo. Sin embargo, si en un momento dado, varias AC que compiten son encendidas, se manifestarán diferentes patrones espacio-temporales y, por lo tanto, la dinámica de la sincronía puede reflejarse en varias bandas de frecuencia. Esta visión ha sido apoyada por conclusiones generalizadas acerca de las oscilaciones y las sincronías en el campo gamma (30-70 Hz) en grupos de neuronas durante las tareas perceptuales. La evidencia experimental ahora incluye registros en varios niveles realizados durante tareas conductuales, desde variadas ubicaciones del cerebro tanto corticales como subcorticales a partir de animales que van de pájaros a seres humanos, y desde señales que abarcan la coherencia de amplitud de banda de las unidades individuales, los potenciales de campo local y los potenciales provocados de superficie (eléctricos y magnéticos). (Ver Singer, 1993; Varela, 1995, para una historia y resumen de estas obras; Lachaux *et al.* , 1998, para resultados recientes).

Esta noción de acoplamiento sincrónico de asambleas neuronales es muy importante para nuestra interpretación de la

temporalidad, y, más adelante, volveremos a ella repetidamente. Este es el punto en que las cosas se vuelven realmente interesantes en nuestro desarrollo de una visión de la cognición, que es verdaderamente dinámica, que utiliza tanto los avances recientes en las matemáticas no lineales como las observaciones neurocientíficas. En resumen, tenemos sucesos constitutivos de nivel neuronal, que tienen una duración en la escala de 1/10, que forman conglomerados que se manifiestan como actos cognitivos incomprensibles, pero completos en la escala del 1. Este período de finalización es dinámicamente dependiente de una serie de asambleas diseminadas y no un período de integración fijo; en otras palabras, es la base del origen de la duración sin un reloj que funciona externa o internamente[14].

Por lo tanto, en esta perspectiva, el ahora es pre-semántico, por cuanto no necesita una rememoración, (o como Husserl dice una "presentificación") para emerger. Una vez más, la evidencia de esta importante conclusión proviene de muchas fuentes. Por ejemplo, los individuos estudiados pueden calcular duraciones de hasta 2-3 seg con bastante precisión, pero su rendimiento disminuye considerablemente por periodos mayores; el discurso espontáneo en muchos idiomas se organiza con silencios que duran de 2 a 3 seg; los movimientos intencionales cortos (como el movimiento autoiniciado de los brazos) se encuentran dentro de ventanas de esta misma duración. Esto trae a primer plano la tercera duración, la escala de 10, propia de las afirmaciones descriptivo-narrativas. De hecho, es totalmente evidente que esta dinámica endógena de los horizontes del ahora puede ser, a su vez, relacionada para formar un horizonte temporal más amplio. Esta escala temporal es inseparable de nuestras afirmaciones descriptivo-narrativas, y se relaciona con nuestras capacidades lingüísticas. Esto constituye el "centro de gravedad narrativo" en la metáfora de Denett (Denett, 1991), el flujo de tiempo relacionado con la

14 El cronometraje preciso es necesariamente flexible (30-100 mseg; 0,5-1,3 seg.) puesto que estos sucesos naturalmente pueden variar en su cronometraje detallado, dependiendo de una serie de factores: contexto, fatiga, tipo de modo sensorial utilizado, edad , etc. Esta es la razón por la cual hablo en un orden de magnitud y no de valor absoluto.

identidad personal (Kirby, 1991). Ésta es la continuidad de un yo que pierde el control por una intoxicación o en patologías como la esquizofrenia o el síndrome de Korsakoff.

Ahora estoy en condiciones de presentar la última idea clave que necesito para completar esta parte de mi análisis: *Los procesos de integración-relajación en la escala 1 son correlatos absolutos de la conciencia del tiempo presente.*

Así, hacemos referencia al ámbito fenoménico y lo que necesitamos estudiar detenidamente es la naturaleza de este vínculo. Las distinciones entre la integración en curso en momentos del ahora, y cómo su integración da lugar a horizontes temporales más extensos en el recuerdo y en la imaginación son el núcleo del análisis husserliano del tiempo íntimo, al que ahora volvemos.

4. El pasado reciente no es recuerdo

Los objetos temporales aparecen ante nosotros como tales, debido sólo a los *actos* correlativos de conciencia que tienen modos específicos de aparición, que se encuentran en el núcleo mismo del problema de la temporalidad inmediata. Normalmente, designamos a estos modos mediante los términos: ahora presente, pasado y futuro. Sin embargo, más allá de esta designación superficial, la reducción apunta claramente al modo del "ahora" como poseedor de un *estatus privilegiado* o único (PZB 35). Dos líneas de análisis conducen a esto. En primer lugar, la textura del ahora, que James denomina "engañosa". En efecto, "ahora" no es una ubicación temporal, puesto que también tiene una cualidad vivida: más que un punto en que un objeto pasa transitoriamente, es un espacio en el cual vivimos. En segundo lugar, se encuentra relacionado con la rica estructura del ahora presente, en que todos los demás modos de temporalidad toman forma.

Por lo tanto, el ahora no es un objeto, sino un ámbito con una estructura análoga al *centro y la periferia* que estructuran el campo visual. El mismo Husserl habla del ahora como un "margen temporal" (PZB 35). En otras palabras, la misma modalidad de

aparición del ahora consiste en una extensión y, hablar de un punto-ahora oculta este hecho: "...presente significa aquí no un simple punto-ahora, sino una objetividad ampliada que, modificada fenoménicamente, tiene su ahora, su antes y después" (PZB 201).

Husserl, aquí, está luchando con lo que Gallagher (1998) llama la paradoja cognitiva de la conciencia temporal: el presente tiene una extensión que es percibida a través de un acto que se encuentra en el presente. Un objeto-suceso temporal, como mi identificación de la figura como una pirámide, tiene una unidad que primero aparece como un ahora presente. Esta, luego, desaparece cuando la figura aparece de nuevo como un vestíbulo. El reconocimiento anterior (y su determinación) se ha hundido en el pasado, como cuando un objeto se mueve en el espacio desde el centro a la periferia. Esto marca el comienzo de la solución de la contradicción aparente entre uniformidad y diferencia, constancia y fluidez, que elaboraremos más adelante. Para hacer eso, ahora debo pasar a otro nivel de detalle en el estudio de la estructura de la conciencia, uno que es constitucional (en la jerga de Husserl), en la medida en que otorga las características temporales de los actos mentales que los unifican en un solo *flujo* de conciencia.

El problema clave en esta fase de mi análisis de la temporalidad es el contraste entre el modo de aparición del ahora y del pasado reciente, el acto que llega más allá del ahora. Como Husserl señala, al comentar un razonamiento similar de Bretano:

> *"No podríamos hablar de una sucesión temporal de tonos si... lo que está primero desapareciera sin dejar rastro y sólo lo que es sentido momentáneamente fuera dado a nuestra percepción" (PZB 397).*

¿Pero, cómo puede constituirse esta estructura de percepción temporal? Lo que se conserva también es modificado. Si cuando viera una pirámide aún pudiera mantener intacto el ahora de cuando vi el vestíbulo, toda la estructura temporal desaparecería. La relación del ahora con el pasado reciente es la de un desplazamiento organizado a través de principios muy estrictos:

"...las presentaciones nuevas, cada una de las cuales reproduce los contenidos de las precedentes, se adhieren a la presentación perceptual y, al hacerlo, se adjuntan al momento continuo del pasado" (PZB 171, las cursivas son mías).

Esta frase expresa las intuiciones que se encuentran detrás del análisis de la *dinámica esencial para estos desplazamientos de la aparición*, tal como se desarrolla en la siguiente sección. Esto es clave para mi búsqueda de puentes entre la naturalización y la textura de la experiencia temporal.

¿Qué forma toma este desplazamiento del ahora al pasado inmediato? El postulado de Bretano era que la constitución del pasado es una re-presentación o una presentificación del recuerdo de lo anterior. En varios comentarios diseminados a lo largo de los años, Husserl realmente se acerca a una demostración de que el desplazamiento del ahora-al-pasado reciente no es igual a la recuperación inmediata del recuerdo o *presentificación*[15]. Para la aparición del ahora-reciente se correlacionan dos modos de comprensión y análisis (es decir, formas válidas de donación en sentido fenomenológico):

(1) recuerdo o memoria evocadora y

(2) imágenes mentales y fantasía.

Existen, al menos, dos argumentos principales para la observación de Husserl. En primer lugar, la naturaleza de recuperación del recuerdo es creada ahora, y seguramente existe un ahora para el acto de recordar. Por lo tanto, no podemos explicar el pasado con un acto que supuestamente sucede en el ahora. En segundo lugar, cuando recuerdo haber visto el vestíbulo (lo que se opone a estar en la situación encarnada viéndolo ahora) el pasado y el pasado sucesivo que se pierde en el olvido tienen una inmediatez, una evidencia para ellos. En el presente, "veo" lo que acaba de pasar; en el recuerdo,

15 Este término poco común traduce bien *Vergegenwärtigung*, una rica mezcla de espera (*Warten*), presente (*Gegen-wart*) y movimiento-desde-el presente (*Ver-gegen-wärttigung*). Esta traducción también marca el contraste con presentación (Gegenwärtigung). Una traducción no literal en un contexto sicológico podría ser memoria evocadora (que incluye el recuerdo y la imaginación), en oposición a memoria icónica o de corto plazo.

sólo puedo conservarlo en una representación como si fuera a través de un velo. Así, el recuerdo y la evocación tienen un modo de aparición que es cualitativamente diferente del ahora[16].

Volvamos a la tarea visual, y resumamos nuestro análisis. Cuando miro la pirámide, me encuentro con el lado cercano en ese momento del ahora y, también, está la unidad del objeto como una unidad duradera (la pirámide es nombrable). Esto revela el juego entre la impresión primaria del lado cercano de la percepción-acción visual y la unidad constituyente que hace aparecer un objeto-suceso identificable. Este es el colector de retenciones, como lo ilustra el ejercicio, que da sentido a todo el suceso: el ahora es experimentado de una forma "original".

> *"Como sólo en un recuerdo primario vemos lo que es el pasado; sólo en él se constituye el pasado, a saber, no de manera representativa, sino, presentativa... La esencia del recuerdo primario es llevar este momento nuevo y único a la intuición primaria directa, así como la esencia de la percepción del ahora es llevar el ahora directamente a la intuición" (PZB 41).*

Por consiguiente, Husserl realiza una distinción disciplinada entre la conciencia impresional opuesta a la conciencia re-*presentacional.* En la impresión, un objeto es originalmente constituido y, por lo tanto, dado como presente (ahora estoy mirando la página y veo la pirámide). En la representación, un objeto-suceso ya dado para la impresión es re-evocado (evoco un poco al ver la pirámide aunque hace rato que la vi por primera vez). Pueden sacarse conclusiones parecidas de la evidencia neurocognitiva. Durante mucho tiempo, la sicología cognitiva ha distinguido entre memoria evocadora, como la insinuada anteriormente, y otras formas de retención inmediata. Por ejemplo, comparar el recuerdo mnemotécnico de elementos rutinario versus el elaborado, produce un cambio sustancial (200-400 mseg)

16 Por lo tanto, la asociación entre las visiones de San Agustín y Husserl acerca del tiempo son equívocas, puesto que el análisis de S. Agustín no distingue entre la presencia del pasado como la memoria evocadora de toda la vida del hombre y el pasado como un presente que se vive.

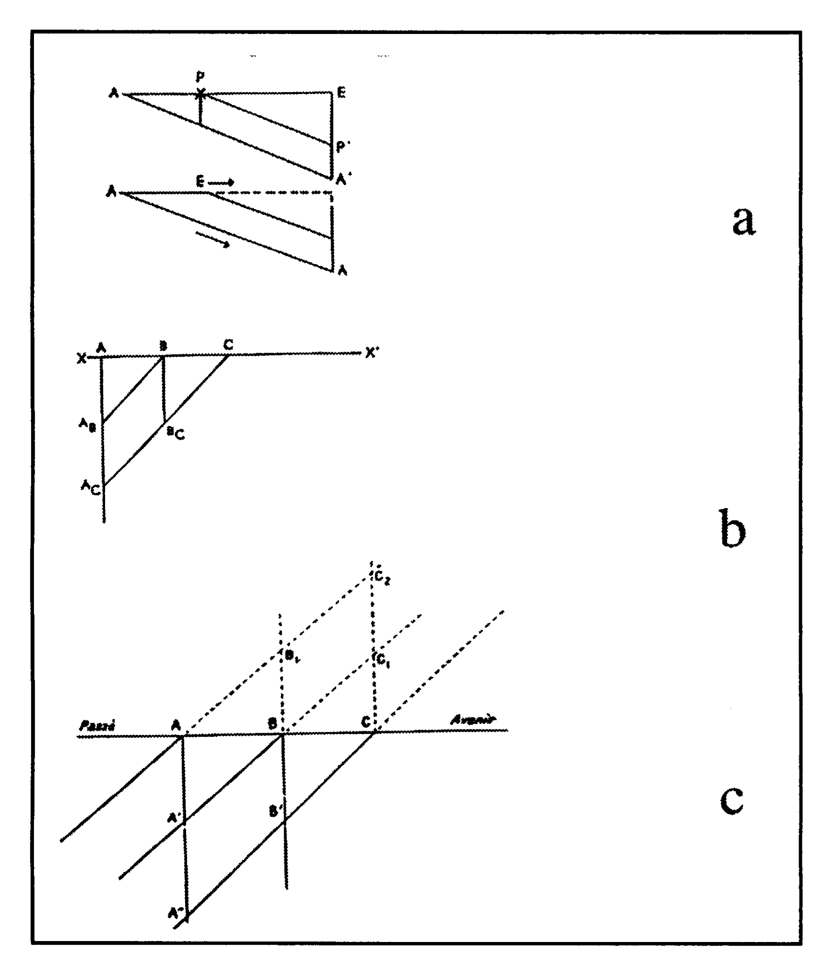


Figura 1. Las “figuras del tiempo”, como las presentó Husserl.

(a) El diagrama elegido por Stein, publicado en las Lecciones (PZB 28)

(b) Una de las muchas variantes seguidas por Husserl antes de 1905.

(c) La versión de Merleau-Ponty (1945, p. 477), que prolonga las líneas en el tiempo de retención, como lo sugirió el mismo Husserl (cf. PZB 331).

en el ERP correspondiente (ver por ejemplo, Rugg, 1995). Más recientemente, los métodos de imágenes cerebrales han comenzado a establecer que durante las tareas presentacionales activas se movilizan estructuras totalmente diferentes, sean éstas un tipo de memoria (episódica, operacional) o imaginación. Todas estas tareas presentacionales movilizan una serie totalmente nueva de capacidades que aumentan el flujo de tiempo, pero que, evidentemente, no se necesitan para esto.

5. La dinámica de la retención

5.1. Las figuras del tiempo: la retención como presente

Husserl introduce los términos *retención* y *protención* para designar la dinámica que sigue la impresión en el presente, que *está dirigida al pasado-reciente* y *al futuro inmediato*. Ésta es la innovación clave en el análisis de Husserl: introducir un nivel de dirección intencional en la conciencia temporal. La retención es el atributo de un acto mental, que conserva un control de las fases del mismo acto perceptual de una manera que se distingue de la experiencia del presente (pero ésta no es una re-presentación, como vimos recién). La característica clave para la retención es su contacto directo con las percepciones anteriores, lo que hace que la percepción, en cualquier momento dado, contenga entidades que se manifiestan como una temporalidad extendida. En la reducción, los objeto-sucesos tienen una duración (son engañosos), pero el acto consciente de la percepción temporal también muestra una temporalidad de tres estratos. Del mismo modo (pero no simétricamente), otra distinción busca tramas futuras o *pro*tenciones.

Esta es la estructura de tres partes que transforma un contenido intencional en una extensión temporal. La figura 1 reproduce tres esquemas para estas "figuras del tiempo" tomadas de fuentes diferentes, pero todas expresan una descripción geométrica de la estructura de tres partes de la temporalidad. El ahora y los niveles de pasado son lo que Husserl llamaba *Ablaufsmodi*, modos de desplazamiento o de transcurso, y que son descritos en referencia a un punto-fuente, de ahí la discontinuidad de este diagrama. La utilización de líneas y

puntos es lamentable, puesto que distrae de la sorprendente agudeza de Husserl; se podría adivinar que esto es un eco de su formación matemática. El destino de estas figuras del tiempo ha sido curioso. Como son fácilmente comprensibles, han sido muy utilizadas como resúmenes de la postura de Husserl. Sin embargo, estas "figuras del tiempo" apenas representan el alcance de sus estudios; en el mejor de los casos ilustran una visión temprana y estática (Fig.1)[17]. El atractivo de los diagramas se hace evidente en la *Fenomenología de la Percepción* de Merleau-Ponty (1945, p. 477), donde el autor comenta con aprobación estas descripciones. No obstante, Merleau-Ponty estaba muy consciente de la deficiencia, puesto que añade cerca del diagrama: "El tiempo no es una línea, sino una red de intencionalidades" (p.477). Y más adelante:

> *"La emergencia de un ahora presente no* provoca *un amontonamiento detrás del pasado y una atracción del futuro. El ahora presente* es *el desplazamiento de un futuro al presente, y del pasado reciente al pasado: es en un solo movimiento que el tiempo se pone en marcha en su totalidad." (p.479).*

Sin embargo, muchos han retomado, sin sentido crítico, este y muchos otros diagramas que aparecen en las obras de Husserl.

En resumen:

> *"la primera impresión está dirigida a la fase nueva y real de los objetos-sucesos temporales, y la presenta como el ahora privilegiado. La retención presenta fases transcurridas recientemente y las presenta en distintos niveles como pasado reciente" (Brough, 1989, p.274).*

La retención es, entonces, un acto específico intencional dirigido al objeto en desplazamiento que lo constituye como pasado reciente. La retención no es un aferrarse al borde del ahora; es una presentación activa de una ausencia que surge de las modificaciones y de la percepción dinámica del ahora. Metafóricamente, se parece más a un movimiento desde el centro hacia la periferia, que a la repercusión de una imagen que se parece al presente modificado sólo

---

17 Para un debate más detallado ver Larrabee (1994) y Miskiewicz (1985).

en intensidad. Pero, además, en el ámbito temporal, ésta es una estructura curiosa: es un pasado de presente-viviente, la retención pertenece a un "presente viviente (*lenbendig Gegenwart*)"[18]. O como Husserl dice, bromeando, en una nota al margen, escrita a mano acerca de sus notas de trabajo: "Pero "pasado percibido" ¿no suena como un *"fierro de madera"? (hölzernes Eisen)* " (PZB 415)[19].

5.2. La retención como trayectorias dinámicas

Es, en efecto, un "fierro de madera" a menos que tomemos una perspectiva dinámica de cómo el origen del ahora puede ser formulado en base a nuestra primera hipótesis. La utilización de los términos "adjuntar" y "desplazamiento" en la descripción fenomenológica ya evoca este punto de vista, pero necesita ser completamente desarrollado. Además del hecho de que existe una base experimental substancial para estas hipótesis, es esencial reconocer que estamos tratando con un auténtico candidato para la síntesis de un espacio temporal donde los sucesos cognitivos se desarrollan.

¿Por qué esto es importante aquí? Porque nos lleva directamente a una visión explícita de los *tipos* particulares de autoorganización que subyacen a la emergencia de las asambleas neuronales. Esto surge de las recopilaciones de una clase particular de oscilador acoplado no lineal, un campo de investigación muy activo (por ejemplo, Mirollo y Strogartz, 1990; Winfree, 1980; Mackey y Glass, 1988; Kelso, 1994). Estos procesos dinámicos, a su vez, aclaran el mecanismo me-

---

18 B. Besnier señala: "Esto es lo que comúnmente se expresa al decir que la retención pertenece a un "presente viviente" (lebenhaftig o *lebending*). Pero hay algo que la retención de-presenta (*démomentanéise*) que puede ser traducido como *ent-gegenwärtigung*, pero esta elección podría crear confusiones..." (Besnier, 1993, p. 339). Para las relaciones con el término Ent-gegenwärtigung como "de-presentación", vea la nota al pie de página anterior acerca de presentación y presentificación.

19 T. van Gelder llamó mi atención respecto de esta reveladora observación, que da el título a su contribución a este tema (van Gelder, 1998). Nuestras ideas respectivas fueron presentadas simultáneamente (aunque desarrolladas de manera independiente) en la Conferencia "Contemporary Issues in Phenomenology and Cognitive Science", en octubre de 1995, en Burdeos.

diante el cual las asambleas neuronales poseen "ahora" una estructura de tres partes.

Los puntos clave para tener presentes en nuestro debate son los siguientes. La autoorganización surge de un nivel *constituyente*, que en nuestro caso ya ha sido identificado como la escala de duración 1/10, y reaparece aquí como un solo oscilador o grupos de osciladores no-lineales. En segundo lugar, necesitamos pensar en cómo estos osciladores entran en sincronía, tal como lo detectó un indicador *colectivo* o variable, en nuestro caso fase relativa. En tercer lugar, necesitamos explicar cómo este nivel de variable colectiva se manifiesta a sí mismo en un nivel global como una *acción* y conducta cognitivas, que en nuestro caso corresponde a la emergencia de un percepto en la multiestabilidad. Este nivel global no es un cálculo abstracto, sino una conducta encarnada sujeta a las condiciones iniciales (por ejemplo, a qué "estaba dirigido", cuál era el percepto anterior) y a parámetros no específicos (por ejemplo, cambios en las condiciones de exposición, modulación atencional). La interdependencia local-global es, por lo tanto, totalmente explícita: la conducta emergente no puede ser comprendida independientemente de los componentes elementales; los componentes alcanzan relevancia a través de su relación con su correlato global.

¿Qué conseguimos? Muy simple. Estos tipos de procesos emergentes nos dan una explicación naturalizada de la discrepancia aparente entre lo que emerge en el presente vivido y la presencia del pasado. En efecto, el hecho de que una asamblea de osciladores acoplados alcance una sincronía pasajera y que le tome cierto tiempo para hacerlo constituye el correlato explícito del origen del ahora. Tal como lo muestran los modelos y los datos, la sincronización es dinámicamente *inestable* y, por lo tanto, dará origen a nuevas asambleas de manera constante y sucesiva. Podemos referirnos a estos saltos continuos como las *trayectorias* del sistema. Cada emergencia se bifurca a partir de las anteriores desde sus condiciones iniciales y límite. Así, la emergencia anterior aún se encuentra presente en la subsiguiente.

Esto trae a primer plano el importante papel de los parámetros de *orden* en las explicaciones dinámicas. Los parámetros de orden pueden ser descritos según dos aspectos principales:

(1) el estado actual de los osciladores y su acoplamiento, o las condiciones iniciales;

(2) las condiciones límite que dan forma a la acción en el nivel global: el marco contextual de la tarea realizada, y las modulaciones independientes que surgen del marco contextual en que la acción se produce (a saber, nuevos estímulos o cambios endógenos en la motivación).

Los parámetros de orden se definen por su encarnación y son únicos para cada caso. Entonces, las trayectorias de esta dinámica envuelven, tanto el surgimiento actual como sus fuentes de origen, en un todo sintético a medida que aparecen fenomenalmente. Se trata en efecto, de una "plancha de madera".

3. La dinámica de la multiestabilidad

Los tipos de dinámica específica en que nos hemos centrado para la comprensión de la retención y del pasado reciente no son simples. Son interesantes, en particular, los sistemas de osciladores acoplados, porque estos, en general, no se comportan de acuerdo con la noción clásica de estabilidad que deriva de una imagen mecánica del mundo. Estabilidad significa aquí, que las condiciones iniciales y límites conducen a trayectorias concentradas en una pequeña región del espacio de la fase en la que el sistema permanece, un punto de atracción o un ciclo límite. En cambio, los sistemas biológicos demuestran que la *in*estabilidad es la base del funcionamiento *normal,* en lugar de una alteración que necesita ser compensada, como en una imagen mecánica de causalidad.

Permítanos volver una vez más a nuestro terreno experiencial de la multiestabilidad visual. El origen de la multiestabilidad se debe a propiedades que son *genéricas* de los osciladores acoplados y sus relaciones de fase. En otras palabras, su modo de aparición es

una invariante bajo ciertas condiciones y cobertura de sujetos. Experimentos recientes, realizados con el fin de estudiar la dinámica de la multiestabilidad en la percepción visual, aclaran esto aún más. Kelso *et al.* (1994) presentaron observadores con *perspectivas divergentes* del clásico cubo de Necker, un pariente cercano de nuestra primera tarea visual. Al pedir al observador que presione un botón cuando se produce el cambio perceptual, se obtiene una serie temporal de los cambios, que obedece a una distribución estocástica (una función gamma con un promedio de 1 seg). Sin embargo, Kelso *et al.* pidió a los observadores realizar la misma tarea de medición de tiempo, mientras anotaban separadamente las sucesiones de tiempo como una función de la perspectiva, que es así utilizada como un parámetro de orden.

Una vez más, es interesante la observación de que en los extremos del cubo la distribución de los intervalos del cambio es considerablemente plana. Es más probable que el individuo experimente cambios esporádicos de la figura, o que permanezca "fijo" en un modo por un mayor tiempo. En estos extremos, los individuos informan ser "bloqueados" por una interpretación. Como antes, uno puede pensar que estos resultados son la forma en que la coordinación de un vasto sistema de osciladores aparece a través de una variable de fase común. Por consiguiente, al introducir variantes de perspectiva, la ubicación en el espacio de la fase es modificada, y aparecen nuevos modos de dinámica que, en este caso, revelan una inestabilidad en forma de silla.

Experimentos recientes demuestran que, tal como suponíamos aquí, esta interpretación dinámica está realmente relacionada con conjuntos de neuronas (Leopold y Logothetis, 1996). Se entrenó rigurosamente a un mono para que "dirigiera" voluntariamente el cambio de un conjunto de figuras ambiguas (rivalidad binocular, una tarea visual que se sabe es parecida al cambio de figuras del cubo de Necker), y luego indicar el momento en el cual este cambio aparecía para su percepción. Al mismo tiempo, se registraron las neuronas individuales de varias de sus áreas corticales visuales. Los autores informaron que en el área sensible al movimiento MT, un

porcentaje de neuronas se correlaciona con el cambio y puede ser modulado por los requisitos perceptuales de las tareas; este porcentaje disminuyó en las regiones principales V1/V2. Este tipo de evidencia apoya fuertemente la idea de que la multiestabilidad surge a través de las colaboraciones a gran escala entre las neuronas, en muchos lugares diferentes en la corteza visual y en otros lugares del cerebro, un ejemplo concreto de un AC emergente para una tarea específica que tiene correlatos perceptuales y fenoménicos.

6. La dinámica del flujo

6.1. El análisis genético de la temporalidad

Necesito abordar, ahora, el último paso de mi análisis, un paso tomado en un terreno mucho menos hollado que la intencionalidad de los objeto-sucesos y la dinámica retencional, dos temas "clásicos" en la fenomenología del tiempo de Husserl. Como el debate ha estado muy concentrado en un tipo particular de acto intencional, la dinámica retencional aún pertenece al análisis constitucional denominado "*estático*", el nivel clásico en el cual ha permanecido la mayoría de los comentaristas del análisis del tiempo de Husserl. Se denomina estático, porque está intencionalmente dirigido a los objeto-sucesos y a lo que aparece (ya sean procesos externos o duración inmanente, por ejemplo, la finalización de un movimiento). En otras palabras, hasta este momento, hemos considerado los niveles más accesibles de temporalidad, la aparición de los objeto-sucesos temporales y los actos de conciencia que los constituyen.

Husserl trae a primer plano un tercer y último nivel de análisis, que ha sido menos estudiado en trabajos posteriores. Este es "el flujo de la conciencia que constituye el tiempo absoluto" (PZB 73). En el contexto de la actual utilización de las ideas dinámicas, es interesante que la elección de Husserl de referirse a este nivel de estudio como "flujo" llame a la reflexión:

> *"Es una subjetividad absoluta y tiene las propiedades absolutas de algo ser denotado metafóricamente como "flujo"* (Fluss*), como un*

punto de la realidad, un punto-fuente principal, del cual nace el "ahora", etcétera" (PZB §36; también t. Nº 54, p.368).

La idea del flujo se abre a dos problemas menos "clásicos" que trataré aquí: (1) el análisis *genético* o *constitucional* del tiempo, presentado anteriormente como el flujo del tiempo absoluto, y (2) su estrecha relación con la dimensión *afectiva* (cf. Sección VII). Husserl estableció la distinción entre análisis estáticos y constitucionales del tiempo durante los años 1917-1923. Estos temas son notoriamente más difíciles de estudiar, no sólo porque se fundan en la última producción de Husserl, sino también porque mencionan áreas delicadas, lo que los hace más atractivos. Por consiguiente, pido al lector considerar lo que propongo en lo que queda de este texto, como un *esquema* de trabajo futuro, más que como cualquier otra cosa[20].

Vuelva a la primera tarea visual (¡Hágalo!). Tenga como objetivo un cambio en el percepto, y vuelva al inicial. Es claro que hemos estado dando dos experiencias distintas con un contenido similar. El vínculo que las une como dos-de-lo-mismo demuestra el hecho básico de que hay una temporalización subyacente, que tiene una independencia relativa del contenido particular de las visiones. Como Husserl señala sucintamente,

> *"Toda experiencia es "conciencia" (*Bewusstsein*) y conciencia es siempre conciencia-de... Toda experiencia es en sí misma experimentada (*selbst erlebt*) y en esta medida, también está dirigida (*bewusst*).* Este *estar dirigido (*Bewussst-sein*) es conciencia de la experiencia (*Erlebnis*)" (PZB 291).*

El vínculo es una reflexión, que no siempre está presente, pero que siempre puede ser puesta en acción, y acompaña todos mis actos. Esta reflexión es temporal, puesto que las experiencias *son*

---

20 Estas dificultades han sido tratadas de manera elocuente por Besnier (1993). De hecho, esta sección le debe mucho a él: me he dejado guiar libremente por algunos de sus argumentos e indicaciones textuales. No he sido capaz de tener en cuenta el material de los manuscritos del tiempo no publicados de Fink (Bruzina, 1993; 1994). El análisis del flujo parece haber surgido en forma estrechamente paralela con los esfuerzos de Husserl por estudiar la intersubjetividad, el problema del tiempo común y la base fenomenológica de la historicidad. Seguramente, esta es una de las áreas más fascinantes y no estudiadas de la investigación fenomenológica. La escacez de fuentes publicadas no hace las cosas más fáciles (ver Depraz, 1992; 1995).

objeto-sucesos (inmanentes) con duración: aparecen deslizándose en el pasado y desaparecen gradualmente en los márgenes del tiempo. La duración de los objeto-sucesos y la experiencia de la temporalidad se constituyen en contra de este trasfondo del flujo de la experiencia. Así, este flujo subyacente da lugar a otra paradoja: éste puede ser separado de los objeto-sucesos temporales, pero al mismo tiempo aparece como inseparable de ellos, puesto que un flujo sin objeto-sucesos no se manifiesta a sí mismo.

La naturaleza de esta inmanencia es dada por Husserl en el siguiente pasaje (notable) que resume su análisis anterior:

> *Puedo expresar la situación de la siguiente manera: Lo que se percibe, lo que se manifiesta (*selbstgegeben ist*) como un objeto individual, siempre es dado en unidad (*Einheit*) con un ámbito absolutamente no manifiesto (*nicht gegeben Mannigfaltigkeit*) (PZB 284).*

La temporalidad inmanente verdadera es la de las experiencias vividas mismas. Aquí llegamos a lo que puede percibirse correctamente como la segunda aporía de la temporalidad: la coexistencia de la permanencia y el cambio. La conciencia es un trasfondo incesante donde aparecen distintos actos y sucesos temporales con su propia duración. Alrededor de 1911, Husserl introduce el nombre de "doble intencionalidad" (PZB 80, 379) para esta articulación, puesto que no sólo existe una retención (del objeto suceso), sino también una *retención de la retención* (una conciencia reflexiva de esta experiencia). Estos dos aspectos de la intencionalidad trabajan unidos y están inseparablemente unidos en el flujo unitario de la conciencia. Puede que la conciencia no exista independientemente de los actos a los que está dirigida o que experimenta, pero sigue siendo distinta de ellos, una unidad llena de aparición y no-aparición.

Ahora podemos preguntar ¿qué vías de acceso o análisis se encuentran disponibles para estudiar el flujo inmanente? Como se dijo antes, una fuente principal es el acto *reflexivo*, el tomar conciencia de la experiencia como temporal. Esto es completamente inmediato, y el argumento más convincente para establecer que este flujo es un fenómeno esencial. Sin embargo, aunque accesible, la reflexión inmediata es una tarea que se hace difícil por el simple hecho de que

para dar una descripción necesitamos continuar con el cambio de la experiencia cambiante. Como hemos dicho, los cambios involucran sólo fracciones de segundo e, incluso, las observaciones minuciosas hacen que sea difícil entregar diferencias útiles[21].

La segunda vía real de acceso al flujo es el recuerdo, si entendemos que dicha presentificación debe ser realizada de manera "pura", es decir, con una visión para su naturaleza y no su contenido específico. Puedo revivir muy claramente el último percepto visual en la tarea. Pero esta evocación sólo es completa cuando también arrastra con ella el contexto encarnado en que la imagen surgió (mi posición, el auto que pasa en el lugar, los fragmentos recurrentes de ideación a medida que estaba haciendo la tarea). En otras palabras, aunque una evocación esté dirigida a un objeto que hace surgir de una manera específica (cf. Sección 4) lo hace como un ámbito: el objeto buscado es un centro, pero también es una periferia llena del contexto de la experiencia encarnada. Este margen, aunque no buscado, es, sin embargo, traído a la vida a través del recuerdo:

> *"Con el fin de estudiar apropiadamente (en su constitución "genética") la temporalidad de la experiencia inmanente, es necesario dirigir el centro reflexivo propio durante una experiencia de* Erinnerung *a las experiencias que son re-producidas por sí mismas. Esto es difícil" (Besnier, 1993, p.374).*

Lo es efectivamente, pero al menos es suficientemente factible que podamos coincidir en que existe algo así como el traer a la vida de los márgenes de un recuerdo. Donde uno puede haber encontrado un objeto aislado buscado por el recuerdo, encontramos, sin específicamente buscarlo, los hilos revividos adjuntos de la experiencia misma. O si puedo hablar metafóricamente: al recordar un objeto buscado, éste sale repleto de los hilos de retención de la

21 Esta es también una afirmación acerca de la pobreza relativa de ejemplos acerca del cultivo continuo de las capacidades reduccionistas, uno de los principales motivos de esta Edición Especial. En el corpus *Abhidharma* del Budismo, se informa que el cultivo a largo plazo de una variante de reducción fenomenológica da acceso una descripción muy detallada de las rápidas sucesiones de la reflexión inmediata (ver Varela *et al.*, 1991 para más información acerca del corpus *Abhidharma*).

experiencia original. Señalado positivamente: lo que hace que el recuerdo tenga este margen retencional debe ser *la forma en que estas mismas retenciones han sido constituidas.*

En nuestro intento de naturalizar los actos retencionales, tratamos de resolver la naturaleza extraña del pasado-siempre-presente. Un análisis más profundo del modo de aparición de la temporalidad inmanente nos enfrenta, otra vez, a una nueva paradoja aparente, que no es distinta de la paradoja del "fierro de madera", que podemos denominar la paradoja *constitucional* de la conciencia temporal. El proceso mismo de desplazamiento (es decir, de-presentación, *Entgegenwärtigung, dé-momentanéisation*) tiene las señales de ser un proceso activo o, incluso, *autogenerado.* ¿Puede existir un proceso que es la causa de sí mismo? En este caso la paradoja toma la forma más clásica de un *regressum ad infinitum.* ¿Existe un ángulo que ilumine esta segunda contradicción aparente?

6.2. La geometría de los flujos no-lineales

Ya se ha esbozado una respuesta, al pasar, y ahora necesita ser completamente desarrollada. La neurodinámica del tiempo que hemos estado siguiendo se basa esencialmente en osciladores no lineales acoplados. Como vimos, este tipo de sistemas dinámicos encuentra su riqueza de conducta en el hecho de que las inestabilidades constitucionales son la norma y no un fastidio que debe ser evitado. El caso de la multiestabilidad hace que, experiencialmente, esto sea bastante evidente: los perceptos saltan de uno a otro por la *misma naturaleza* de la geometría del espacio de la fase y las trayectorias.

Esta es una caracterización generalizada, no sólo aplicable a este estudio de caso. En general, los sistemas complejos, no lineales o caóticos otorgan un automovimiento que no depende (dentro de una gama de parámetros) de donde están los sistemas. En otras palabras, si el contenido de mi percepto visual es un hombre/mujer o una pirámide/vestíbulo, el movimiento intrínseco o inmanente es *genéricamente* el mismo. Si el lugar específico en la fase espacial es un correlato del contenido intencional de un objeto-suceso, el

sistema nunca se detiene demasiado en ello, sino que se acerca, toca y se escabulle en un movimiento perpetuo y autopropulsado. Cognitivamente, esto corresponde a la observación de que en el cerebro y la conducta nunca existe un estado cognitivo que se detiene o mora, sino un cambio permanente interrumpido por los conglomerados transitorios que subyacen tras un acto momentáneo (escala de duración 1). Esto se expresa formalmente a través de la presencia dominante de regiones estables/inestables, de manera que cualquier cambio leve en las condiciones iniciales y límites hace que el sistema se mueva hacia una región cercana estable/inestable.

Esta idea ha sido, y aún es estudiada por la dinámica de varias maneras diferentes. En el caso específico de la multiestabilidad del cubo de Necker, asumimos la coordinación de un amplio sistema de osciladores a través de una variable de fase común. Mediante la introducción de variantes de perspectiva, el paisaje dinámico es, por consiguiente, cambiado por nuevos modos dinámicos que aparecen en el retrato de la fase. Esta inestabilidad en forma de silla supone que en este punto hay una mezcla de tendencias para ser atraídas a esa posición (es decir, el contenido perceptual) o para alejarse de ella, y la más leve perturbación empujará y tirará de las trayectorias.

De manera más general, muchos sucesos naturales aparentemente "ruidosos" (como los períodos de transición entre los cambios del cubo de Necker) han arrojado, recientemente, patrones deterministas no esperados en un análisis dinámico no lineal, más allá del alcance del análisis lineal tradicional. La principal característica de estos métodos es darnos una visión de la dinámica no sólo basada en trayectorias, sino en el marco más amplio de la *geometría del paisaje del espacio de la fase*.

Esta no es sólo una descripción formal. Estos patrones geométricos pueden ser estudiados incluso en una agrupación de actividad neuronal muy localizada, como la medida por un electrodo en la superficie del cerebro (unos pocos milímetros cúbicos de la corteza). Por ejemplo, en un centro epiléptico temporal local, que se parece a una oscilación ruidosa, hemos encontrado evidencia para múltiple determinismo e inestabilidades (Le van Quyen *et al.*, 1997

a, b). Mientras grabábamos los registros del cerebro de un individuo, se le pidió que realizara una discriminación visual y auditiva simple. Entonces, estudiamos el intervalo entre estas descargas, de manera similar a las sucesiones temporales del cambio de percepto de Necker. Esto dio indicios sistemáticos de que estas dinámicas temporales no pueden ser caracterizadas como una simple periodicidad "ruidosa". Incluso un simple mapa de primer retorno (el valor de un intervalo de tiempo determinado contra el valor del siguiente), revela patrones geométricos cambiantes detallados que dependen de las condiciones experimentales, pero muestra evidencia sistemática de una inestabilidad en forma de silla con colectores estables e inestables. Esta observación sugiere que el paisaje del espacio de fase puede caracterizarse por *desviaciones de la estricta periodicidad en una forma no-aleatoria* (Auerbach *et al.*, 1987). En particular, las diferentes discriminaciones perceptuales "arrastran" esta dinámica local hacia una distinta periodicidad inestable. Aunque estas posiciones de los puntos periódicos sean traspasadas entre las condiciones conductuales estudiadas, las pendientes relacionadas del enfoque para las inestabilidades aparecen como características invariantes de la dinámica para todas nuestras condiciones experimentales.

En este estudio de caso, enfatizamos la importancia de las propiedades locales no lineales en los sucesos cerebrales, que, a menudo, se pierden cuando se aplican métodos globales promedio (como los estimadores de espectro o, incluso los de dimensión). Existe un sorprendente nivel de detalle mostrado por las trayectorias u órbitas de la dinámica epiléptica a medida que es modulada por las tareas perceptuales. Una manifestación paradigmática de esto es el hecho de que una trayectoria caótica incluye típicamente una cantidad i*nfinita* de órbitas periódicas inestables. Estas órbitas son inestables en el sentido de que la más mínima desviación saca al estado de su órbita periódica. Por lo tanto, un sistema determinista no lineal (o caótico) nunca permanece por mucho tiempo en ninguno de estos movimientos inestables, sino que cambia continuamente de un movimiento periódico a otro, dando así, una apariencia de aleatoriedad (Artuso *et al.*, 1990).

6.3. La doble intencionalidad

Hemos obtenido, entonces, una intuición renovada para resolver el enigma de este segundo elemento extraño de la mezcla de pasividad y actividad, de invariancia y cambio de la doble intencionalidad. "La autoaparición (*Selbsterscheinung*) del flujo no necesita un segundo flujo, más bien, se constituye a sí misma como un fenómeno en sí (*in "sich"selbst*)" (PZB 381). Merleau-Ponty se refiere a este paradójico aspecto de la descripción reductiva diciendo que para existir, el tiempo "ya debe estar en mí (*fuse en moi*)", y, a medida que éste surge como un flujo de retenciones, también se automanifiesta (*Selbsterscheinung*)[22]. De hecho: "No elegí nacer, pero una vez nacido, el tiempo debe permearme (*le temps fuse à travers moi*), haga lo que haga" (Merleau-Ponty, 1945, p. 488).

Husserl desarrolla su explicación descriptiva de esta aparición paradójica de "doble intencionalidad" en las nociones de intencionalidad transversal y longitudinal (*Quer-* y *Längsintentionalität*)[23]. La primera, es la dinámica retencional, la constitución estática. La intencionalidad longitudinal, en cambio, es la constitución genética de la temporalización de las experiencias mismas, su automanifestación. Éstas son necesariamente interdependientes, pero su modo de dependencia y la raíz de su diferencia son difíciles de expresar: "No tenemos nombres para todo eso" (PZB 371).

Ahora, la intencionalidad longitudinal actúa mediante una integración desde el interior del mismo ahora, lo que otorga un sustrato inalterable del cual emerge el flujo. Como señala Besnier (p.350), existe una gran tentación de transponer el análisis de la intencionalidad perceptual concibiéndola como un sustrato "puro" o *Ur-hyle*. No tengo que entrar en el peliagudo debate técnico que este tipo de análisis ha producido desde Husserl (Depraz, 1996). Mi contribución es traer a primer plano la intuición

22 Merleau-Ponty (1964), p. 244-5, 296-8, como lo trató Besnier, 1993, nota 9 al pie de página, p. 356.

23 Cf. PZB §39, t. Nº.54.

derivada de los flujos genéricos no lineales. La automanifestación aparece en nuestro análisis como un automovimiento o inestabilidad genérica, que no es un simple artefacto o descripción, sino una descripción formal invariante para la autoorganización. Por lo tanto, su importancia para la temporalidad es apropiada. El flujo, en el sentido neurodinámico, es precisamente un fenómeno extraño, que existe como un flujo sólo en la medida en que es constituido en trayectorias individuales (no un magma geométrico inerte), a medida que las trayectorias momentáneas autopropulsadas visitan varias regiones en el espacio de la fase (que corresponden a un objeto-suceso buscado, una aparición).

Aquí, la inseparabilidad de estas dos intencionalidades no sólo es descriptivamente exacta, sino parte de la lógica intrínseca de dinámica no lineal compleja. Sería contradictorio calificar el automovimiento como un "estrato profundo" del proceso dinámico, y describir las trayectorias como mera apariencia (Gallagher, 1979). *Mutatis mutandis*, parece ilusorio aislar un estrato "más profundo" de una constitución genética, en que la experiencia sería constituida a partir de un tiempo absoluto, y sólo entonces se manifestaría en la intencionalidad consciente. Lo que es profundo es el *vínculo* entre el automovimiento (inmanencia) y las trayectorias (aparición).

Se ha dicho suficiente acerca del flujo inmanente o absoluto para sugerir su importancia y perspicacia. Es, sin duda, un tema que necesita ser estudiado, y es el fundamento natural para los puentes en otras variedades de experiencia y hacia otras tradiciones preocupadas de la conciencia humana. Además, este nivel de análisis toca más que cualquier otro el terreno del yo, del ego puro, o conciencia básica. Brough (1989) resume:

> *"Y gracias al horizonte infinito abierto por el flujo absoluto, podemos estar seguros de que podemos seguir cambiando y acumulando un pasado correctamente, aunque siga siendo el mismo. Entonces, hay una fisura en la conciencia. Gracias a esta fisura, cada uno de mis actos, que se reafirman a sí mismos y se hacen valer, cambia por un tiempo más o menos breve... es capaz*

de "deslizarse" sin tomar todo mi ser con él. Si mi conciencia interna estuviera pegada sin interrupción a mis breves experiencias, el paso del tiempo destrozaría mi ego" (p. 288).

Sin embargo, ahora debemos dejar estas consideraciones y volver a nuestro último tema: el tema estrechamente relacionado de la aparición de este automovimiento desde la perspectiva del afecto.

7. Protención — Transparencia y tono emocional

7.1. La temporalidad inmanente y el afecto

Volvamos de nuevo a la primera tarea y re-examinemos más detenidamente la naturaleza del cambio, tal como sucede. Un componente esencial de la experiencia es que el cambio es repentino y va acompañado de un cambio emocional (más o menos distinto) cuando la percepción visual cambia abruptamente. De esta manera, el modo de desplazamiento del ahora en el pasado reciente, la trayectoria retencional, aparece como la presencia del pasado, no sólo en una manera que es distinta de la memoria representacional. Esto también nos da la clave de que el *tono emocional* es una parte integral del fenómeno. ¿Cuál es el papel de la emoción o del afecto en el automovimiento del flujo? ¿Y cuál es su papel, si tiene alguno, en la anticipación de lo que vendrá, la protención?

En los textos publicados de Husserl, la protención no es exhaustivamente analizada y tengo la impresión de que él supone implícitamente cierta simetría con la retención, como si la misma estructura de invariancia para el pasado pudiera ser lanzada hacia el futuro. Pero la protención dirige lo nuevo anterior hacia una impresión y, así, sólo puede ser una pre-figuración. Husserl habla de una "constitución vacía" (PZB 52), pero no es una expectativa o anticipación en el sentido de que contiene una representación de lo que el próximo ahora traerá. Para ver por qué esto es así, uno sólo necesita aplicar los mismos argumentos utilizados para distinguir la retención y la memoria evocadora. En efecto, la protención tiene un modo de apertura: "...la única cosa definitiva es que, sin excepción, algo sucederá"; al escuchar una melodía (su ejemplo) hay un aspecto

predecible hacia la protención, puesto que está dirigido a más frases de la música (PZB 106, 84). Estas y otras observaciones similares de las *Lecciones* parecen haber dado el fundamento para la visión simétrica de la estructura de tres partes del tiempo y, como se indica en la muy utilizada figura del tiempo tratada anteriormente (Fig.1).

Este análisis puede enriquecerse sustancialmente. Existen al menos dos fuentes principales de evidencia para concluir que la protención es genéricamente n*o* simétrica con la retención. Lo primero es, precisamente, que lo nuevo es siempre invadido por el afecto y el tono emocional que acompañan al flujo. De hecho, la protención no es un tipo de expectativa que puede ser comprendida como "predecible", sino una apertura que es capaz de realizar un automovimiento, una indeterminación que está a punto de manifestarse. En esta cualidad, la protención proporciona el vínculo con el afectar o, más acertadamente, con alguna forma de autoafectación (ver más adelante). Lo segundo, es que la retención tiene la estructura de lo continuo, pero la protención sólo puede ser un ámbito limitado, puesto que no puede prever una anticipación que aún está por venir. Mientras los hilos de la retención crean el marco para la protención, ésta no puede modificar retroactivamente los hilos retencionales[24].

El tiempo y el afecto nunca fueron tratados sistemáticamente por Husserl. Una parte importante de sus notas aún no es publicada o sólo es accesible a partir de fuentes secundarias. Sin embargo, una profundización de los últimos análisis del tiempo hace posible descubrir algunos fragmentos de una visión del afecto mientras inicia el camino del flujo vivido mismo (Depraz, 1993). Así, mis proposiciones en lo que queda de esta Sección no son completamente divergentes con la última investigación de Husserl, pero vamos más allá de esto. Como él dice:

> *¿De qué manera el yo (*Ich*) es el centro de esta vida que experimenta? ¿Cómo es experimentada? Es afectado por aquello de lo que la conciencia está consciente (*Buwusstsein bewusst ist*), sigue*

24 Gallagher (1979) también trata esta asimetría, y señala que Husserl habla de las diferencias de "estilo" entre protención y retención (*Hua. XI*, p. 323-4).

el afecto, o aún, es atraído, sujetado, asimilado por lo que lo afecta (Ms. CIII/1).

Las notas de Husserl contienen alusiones recurrentes a este aspecto primordial relacionado con la temprana vida de los niños en que se encuentra una intencionalidad "instintiva" (*Triebsintentionalität*). Como señala Depraz:

"El afecto es entonces esta no-forma, que realiza la constitución del yo por sí mismo, que lo afecta en el estricto sentido de la estructura, la de la temporalidad constitutiva... El afecto está allí antes de estar allí para mí en conciencia total: soy afectado antes de saberlo. Es en este sentido que se puede decir que el afecto es primordial" (Depraz, 1993, p. 73,75).

¿De qué manera esto es pertinente para la estructura de tres partes del tiempo? Husserl señala que durante una melodía, el sonido me afecta de manera distinta a medida que crea sus hilos retencionales, una tendencia atencional (*eine Tendez der Zwendung*). O podemos decir que ésta otorga una *disposición* que está marcada por unas intensidades graduales. Este efecto temporalizador pone a la protención en el escenario central: "... no es sólo la impresión de la materia (hyle) lo que afecta, sino las "anticipaciones hyléticas de los datos" " (*ibíd.*, p. 79).

El tiempo siguió siendo después de Husserl un camino real hacia el estudio del afecto, especialmente con Heidegger y Merleau-Ponty, en que el debate acerca de la autoafectación es considerablemente enriquecido. La principal innovación radica en su tratamiento del tiempo como autoafectación. Como dice Merleau-Ponty: "El tiempo es afección del yo por el yo", y señala que la expresión deriva de Kant y fue modificada por Heidegger en su *Kantbuch*[25]. La autoafectación se transforma en una perspectiva clave en la naturaleza de la conciencia:

"...incluso la conciencia más precisa de lo que somos capaces es afectada por sí misma o dada a sí misma. La misma palabra

25 Merleau-Ponty, 1945, p. 487; "*Die Zeit ist ihrem Wesem nach reine Affektion ihrer selbst*", *Kant und das Problem der Metaphysik*, p. 180-1.

conciencia no tiene significado aparte de esta dualidad." (Merleau-Ponty, 1945, p. 488).

Con Levinas (1988), se trae a primer plano otra esfera de la afección, la hetero-afección: El Otro, la alteridad, es la clave principal para la constitución del tiempo. No sólo somos afectados por representaciones y afecciones inmanentes ("afection de soi par soi"), sino también por la alteridad, como inseparable de la esfera de un yo-ego. En esta medida, la misma distinción entre auto- y hetero- deja de ser importante, puesto que en todos los casos todo se reduce a la misma manifestación: es una cuestión de "algo otro", la experiencia de una alteridad, una diferencia en la identidad del presente, ya sea a través de los inevitables desplazamientos hacia la retención o, a través de las anticipaciones de la protención.

Pero, aquí, estos refinamientos filosóficos no son centrales para nosotros en todo su detalle. Buscamos ir más allá de la paradoja aparente entre una impresión original en el tiempo, que sería coloreada por la afección, o, a la inversa, la primacía de la afección que subyacería a la temporalidad. Buscamos una síntesis no dual en que el afecto sea constitutivo del yo, y al mismo tiempo contenga una apertura o imprevisión radical en relación con su ocurrencia.

7.2. Disposición para la acción

Con el fin de avanzar más en nuestro análisis, necesitamos tener una base concreta sobre la cual fundamentar nuestro estudio, así como las tareas visuales daban la base para el análisis estático. Existen muchas alternativas, pero he elegido estudiar el papel de la afección en la constitución del tiempo, en el contexto de la participación activa en el mundo y a través de la calidad disposicional del afectar y sus gradaciones, puesto que —desde el punto de vista de la neurociencia cognitiva enactiva — el enfrentarse con la vida juega un papel central. El enfrentarse con la vida es un estar preparados o una tendencia disposicional para la acción en un ámbito más amplio, una *disposición ontológica*, es decir, una expectativa de la manera en que el mundo se pondrá de manifiesto. Por esta misma razón, enfrentar algo tiene mucho que ver con el hábito, la repetición de

nuestras vidas. Aprender una habilidad es un ejemplo prototípico de la adquisición de transparencia para la acción.

Como hemos visto, la pérdida de fluidez en el enfrentar la vida nunca está lejos de un tono-afectivo. Pero ahora podemos ver diferentes grados de quiebre en la transparencia y las múltiples formas en que esto sucede abren una colección de tonalidades afectivas: miedo, celos, rabia, ansiedad, seguridad de sí mismo, etc. En consecuencia, la palabra emociones es utilizada aquí en su sentido específico: la *tonalidad del afecto que acompaña un cambio en la transparencia*. Por otro lado, el afecto es una ampliación de la orientación disposicional que precondicionará el tono emocional que puede aparecer.

Ahora, a medida que escribo, tengo una actitud disposicional que me compromete en una expectativa de escribir y dar forma de oraciones a mis pensamientos. Cuando escribo esta palabra ahora, la disposición es coloreada por una carga emocional, un resentimiento moderado por no encontrar la expresión apropiada. Sin embargo, este tono emocional aparece contra el telón de fondo del ánimo exaltado de un día productivo dedicado a terminar este texto.

Para ser más explícito, quiero distinguir tres escalas para el afecto, homólogas (pero no isomórficas) a las tres escalas de temporalidad utilizadas anteriormente.

(1) La primera escala son las em*ociones*: la conciencia de un cambio de tono que es constitutivo del presente vivido.

(2) El segundo es el *afecto*, una tendencia disposicional propia de una secuencia coherente de acciones encarnadas.

(3) Finalmente el *ánimo*, la escala de la descripción narrativa de una duración más o menos larga.

7.3. El tono emocional como paisaje dinámico

Para el etólogo, el afecto y las emociones son un repertorio relativamente pequeño de disposiciones inmediatas que están inscritas fisiológicamente en la herencia de las especies, aunque en la mayoría

de los mamíferos el hábito y el aprendizaje continuo pueden formarlo significativamente[26]. Neurobiológicamente, estos pueden ser asociados con un conjunto relativamente estable de correlatos neuronales (por ejemplo, Damasio, 1994). Los estudios de las respuestas emocionales humanas, incluso en situaciones relativamente artificiales, revelan la medida en que el legado biológico de patrones emocionales básicos está envuelto en la repetición histórica de un individuo, su historicidad y lenguaje. Los hábitos individuales, la historicidad y el lenguaje constituyen la gama de la vida emocional humana, que incorpora la estructura biológica a un final que es, pero que es histórica e individualmente único[27].

Homólogamente, podemos decir que la experiencia del tiempo tiene una base biológica en los sucesos elementales (escala 1/10), pero esta base está envuelta con otras estructuras de temporalización en el presente engañoso que es nuestro tema. Es infructuoso negar que esta base biológica profundamente arraigada tiene un papel en la aparición de la temporalidad. Asimismo, no estoy reduciendo las emociones a su correlato empírico en un paso reduccionista. Las emociones, tal como se consideran aquí, son una parte integrante de un estar preparado. Esto no debería ocultar el hecho de que esta constitución ontológica tiene su origen en las disposiciones emocionales básicas inseparables de nuestra historia como seres vivos y de los mínimos sucesos en la fisiología del cerebro.

Cuando provoco una interrupción de los hábitos, al ejercer la reducción cuando miro la imagen visual de nuestra tarea, le llevo una disposición emocional que prefigura el cambio en mi percepción. Al decir, "espero ver" también entrego parámetros de orden adicionales y exógenos, que alteran la geometría del espacio de la fase. Este proceso de "esculpir" un paisaje dinámico es

26 Para los estudios básicos de instinto y emoción en los animales ver los estudios fundamentales de K. Lorenz (1972), que aún son útiles. Para un estudio más reciente y que incita a la reflexión ver der Waals (1995).

27 Ver Ekman y Davidson (1994) para un estudio reciente de estudios sicológicos, y Rorty (1980) para las visiones cognitivas.

intrínsecamente distinto de las trayectorias que se mueven dentro de él, pero forman una unidad inseparable.

De hecho, se ha sabido, por algún tiempo, que la intención de realizar un movimiento está acoplada con un cambio en el tono emocional que varía en nivel. Como variable global, los cambios inducidos en el paisaje dinámico pueden ser detectados. Un caso bien conocido es el potencial de estar preparado. Para el movimiento de un dedo, se puede medir un gran potencial eléctrico lento en todo el cuero cabelludo, que pr*ecede* por una fracción de segundo el comienzo del movimiento, y el individuo puede informar que ha decidido iniciar el movimiento[28] . Este no es un correlato de la intención (como a veces se dice), sino que da una idea concreta de cuán extensamente se involucra una reconfiguración de un paisaje dinámico en el origen de un ahora completamente constituido (el movimiento de los dedos). En los resultados de Leopold y Logothetis (1996), ya mencionados, una reconfiguración similar de la disposición para el encendido de las respuestas neuronales es visible unos 100-200 mseg *antes de* que el mono indique que ha cambiado a un nuevo percepto.

¿Por qué todo esto es importante aquí? Porque es una evidencia directa de la manera en que la tonalidad emocional actúa en la dinámica de flujo. La tonalidad emocional es, por su misma acción, un límite importante y una condición inicial para la neurodinámica. Este difuso efecto constitutivo está de acuerdo con el mecanismo de acción a través de los neurotransmisores, que se sabe condicionan las modalidades de respuesta a nivel neuronal, tal como lo atestigua la corriente de conocimiento de los agentes psico-farmacológicos.

Este esquema de la naturaleza de las protenciones a través de la tonalidad-afectiva nos ha llevado a un tercer y último paso de lo que parece ser formalmente una constitución genética de la temporalidad. He introducido un último principio dinámico que

---

28 La descripción original se debe a Deecke *et al.* (1969). Desde entonces se han publicado muchos trabajos acerca de estos correlatos eléctricos pre-concientes.

se aplica también a la dinámica neurocognitiva. Me refiero a la circularidad mutua entre el paisaje del espacio de la fase y las trayectorias específicas que se mueven en él, y el hecho de que las mismas trayectorias entregan las mismas condiciones para un acoplamiento encarnado, puesto que a través de su acoplamiento forman su paisaje dinámico. Metafóricamente, el caminante y el sendero están intrínsecamente vinculados.

Este principio de circularidad parece estar presente en una variedad de sistemas naturales, y recientemente se ha dicho de él que "funciona al borde del caos", o que está "autoorganizado críticamente". Por ejemplo, esta idea da una visión renovada de la evolución, puesto que brinda una respuesta al antiguo dilema de naturaleza (expresión genética) versus crianza (condiciones medioambientales de acoplamiento). En esta visión sintética (cf. Kauffman, 1993), la relación entre las formas naturales, (el *Baupläne* de los organismos), y el proceso de selección en su incrustación ecológica, no es de contradicción, sino precisamente una imbricación mutua cuando se ve a través del lente de la dinámica[29]. Esta movilidad inherente, que envuelve trayectorias y geometría, da a un sistema natural la posibilidad de estar siempre cerca de regiones en el espacio de la fase que tengan múltiples recursos (por ejemplo, al menos dos en la percepción visual bi-estable).

Para terminar: La estructura genérica de doble intencionalidad propuesta por Husserl es, lo admito, de este tipo de enlace dinámico y el análisis fenoménico del afecto y la tonalidad emocional proporciona evidencia de ello.

8. El ahora — Nuevas figuras del tiempo

Para atar todos los hilos que he estado desarrollando aquí y hacer eco de la tradición comenzada por el mismo Husserl, me gustaría proponer una nueva figura del tiempo, la *estructura cuádruple del ahora* (ver más adelante, Figura 2).

29 Esto se refiere a la evolución en el sentido natural en Varela *et al.* , 1991.

Esto no es tan atrevido o exagerado. Por ahora, parece claro que no es suficiente la descripción del tiempo lineal, punto por punto en la base de las figuras del tiempo. Una mejora importante es introducir, no puntos y líneas, sino flujos, tendencias *dinámicas.*

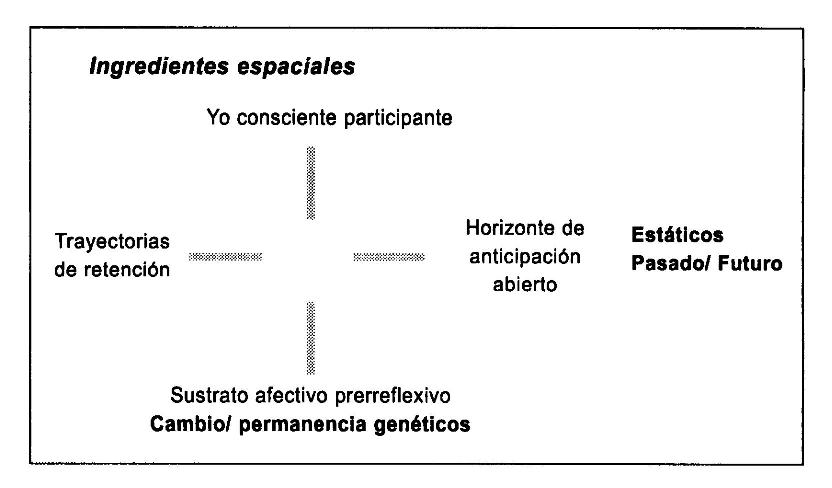
Una segunda mejora importante, es tomar como explicación explícita lo que surgió en el último trabajo del mismo Husserl, el papel central de la doble intencionalidad, la constitución estática y genética. Este ingrediente final da a las homologías entre la constitución del espacio y los tiempos la preeminencia que merecen. Hago esto tomando la estructura de *centro/margen* como el mismo núcleo de una nueva figura del tiempo. Una vez que estos tres aspectos básicos hayan sido incorporados, una nueva representación se aclara de manera bastante natural.

1. Para comenzar permítanos considerar el papel de la dinámica, que ha sido central en nuestro desarrollo, y de ahí, pasar a la utilización de los puntos discretos en línea como en el diagrama tradicional. No sólo dejamos atrás una figura geométrica lineal, sino que también debemos introducir una *asimetría* en ella. En relación con la constitución estática, hemos discutido dos clases diferentes de ideas dinámicas: las trayectorias retencionales (el pasado) y los parámetros de orden para la anticipación (el futuro). En el nuevo diagrama de flechas en hileras, como es tradicional en las matemáticas, se muestra la cualidad dinámica, pero distinguí las trayectorias del paisaje anticipativo haciendo descender las cabezas de las flechas. Con respecto a la constitución genética, también tenemos dos ideas diferentes: la temporalización inmanente del automovimiento y la intencionalidad dirigida relacionada con una posición en el espacio de la fase.

Ingredientes dinámicos

trayectorias de retención	parámetros de orden de anticipación	**Estáticos** pasado/futuro
atrayentes transitorios contenido intencional	automovimiento de las inestabilidades genéricas	**Genéticos** cambio/ permanencia

2. A continuación consideramos los ingredientes espaciales, es decir, el papel de una configuración centro-periferia en el núcleo de la temporalización. En lo referente a la constitución estática, por supuesto, recuperamos el eje retención-protención. Una vez más, estos son asimétricos, puesto que las retenciones se pierden en el pasado, pero el margen protencional es un horizonte de anticipación abierto. Así, descubrimos, una vez más, que la estructura centro-periferia no entrega todas las distinciones necesarias, puesto que a medida que los márgenes se alejan del centro, se vuelven cualitativamente diferentes. Como en la constitución genética, en una dirección reaparece el margen en el sustrato preconsciente, afectivo (permanencia), y en la otra, el ego consciente, encarnado y consciente del cambio emocional (cambio).



Así llegamos a ingredientes que vienen es d*os series de cuatro*, cuatro aspectos sugestivos. Las configuraciones de centro/periferia son análogas en los esquemas espaciales y dinámicos, puesto que ponen de relieve la naturaleza paradójica del pasado en el presente, y del cambio en la permanencia que desconcertaron a Husserl (y a muchos otros) a lo largo de su trabajo. Estos pares son los elementos extraños de la figura del tiempo, que aparece aquí bajo una luz nueva y no paradójica.

Fundamentalmente, la perspectiva adjunta proviene del hecho de que los ingredientes constituyentes tienen un vínculo *genérico* entre ellos, una interdependencia interna que ha sido estudiada a lo largo de este artículo. En otras palabras, las nuevas figuras del tiempo no son sólo combinaciones gráficas de puntos, sino que muestran vínculos efectivos, que no son sólo descriptivos. Fenomenológicamente, he enfatizado la total *interdependencia* de ambas intencionalidades, la inseparabilidad del análisis estático y genético y la determinación mutua de la constitución instintiva y cognitiva del yo. En forma paralela, las trayectorias y los paisajes de su espacio de fase son una unidad en un sistema complejo y no lineal.

En consecuencia, hemos estudiado los muchos aspectos según los cuales el determinismo, las trayectorias, las regiones en el espacio de la fase, los paisajes geométricos adaptivos son complementarios. Considero a estas interdependencias mutuas y su papel en la constitución de la temporalidad como las perspectivas más inmediatas que la naturalización puede entregar.

Tal como dije, con estos elementos en situación, la figura del tiempo naturalmente se aclara. La figura 2 coloca a estos ingredientes en relación con los demás, con respecto al papel de la dinámica y el espacio y la causalidad circular para la intencionalidad longitudinal, pero no para la intencionalidad transversal. Como los parámetros de orden para la protención no influyen retroactivamente la historia del sistema, llegamos a una figura asimétrica cuádruple del tiempo.

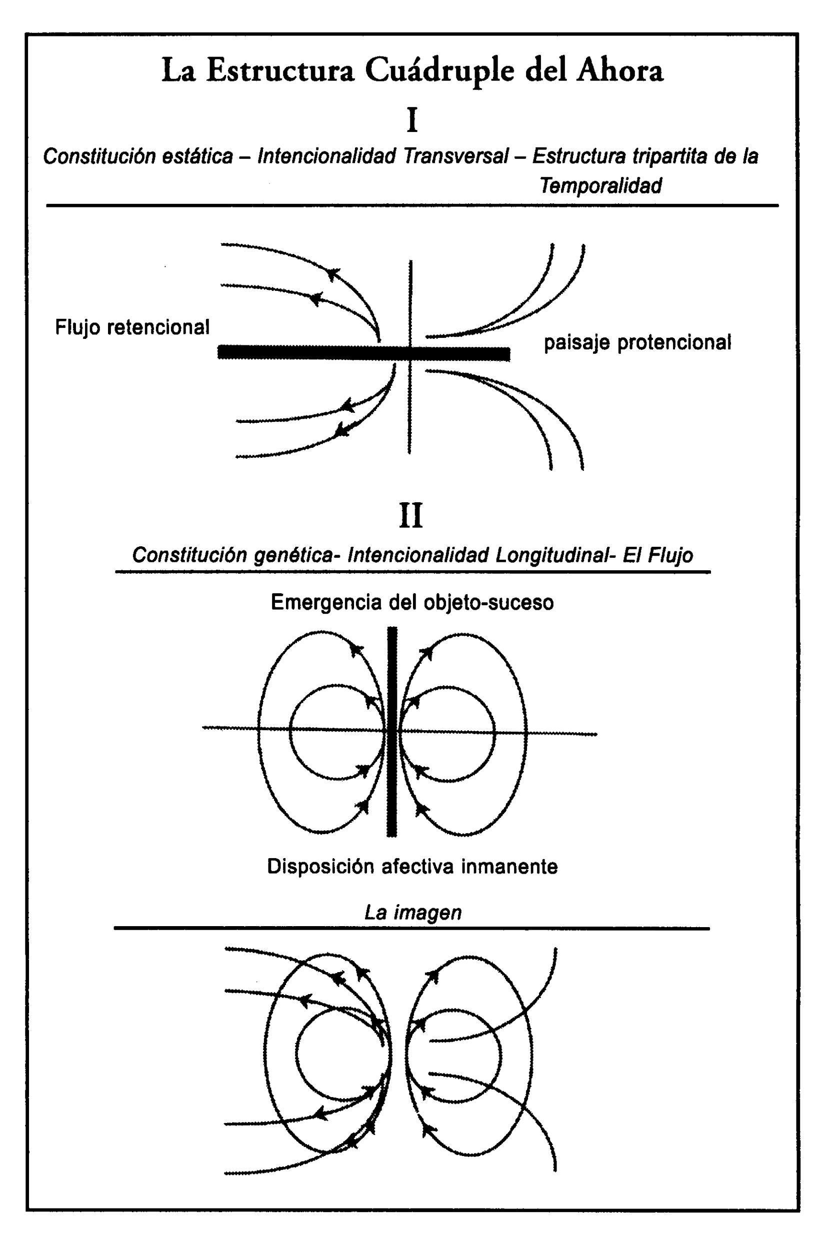


Figura 2.

Las nuevas figuras del tiempo resultantes del análisis neurofenomenológico presentado aquí. Para mayores explicaciones, véase la Sección 7.

9. Colofón — La trenza triple de la neurofenomenología

Ahora estamos en situación de distanciarnos de nuestro análisis de la temporalidad y reconsiderarlo. ¿La neurofenomenología ha pasado la prueba mordaz? Permítannos volver al programa de investigación neurofenomenológica: la circulación entre lo externo y lo experiencial. En Varela (1996), la hipótesis de trabajo neurofenomenológica fue planteada de la siguiente manera:

> Las explicaciones fenomenológicas de la estructura de la experiencia y sus equivalentes en la ciencia cognitiva se relacionan entre sí a través de limitaciones recíprocas.

Por una parte, estamos interesados en un proceso de emergencia externa con atributos neurobiológicos bien definidos; por otra, en una descripción fenomenológica que es cercana a nuestra experiencia vivida. La naturaleza de la circulación que se busca no es menor que las *limitaciones mutuas* entre ambas explicaciones, incluidos los puentes potenciales y las contradicciones entre ellas. ¿Cuál es la naturaleza específica de los pasajes entre estas dos explicaciones en el presente caso? ¿Qué hemos aprendido en el caso específico de la temporalidad inmediata?

Una cosa es clara: la naturaleza específica de las limitaciones mutuas está lejos de una simple correspondencia empírica o un isomorfismo categórico. Tres ingredientes han resultado tener un papel igualmente importante: (1) la base neuro-biológica, (2) los instrumentos descriptivos formales en su mayoría derivados de la dinámica no lineal, y (3) la naturaleza de la experiencia temporal vivida estudiada según la reducción. Lo que necesita ser estudiado detenidamente es la manera en que esta trenza triple, estos tres ingredientes están *trenzados en conjunto de manera constitutiva* (ver Varela, 1997 para más información acerca de esta neurofenomenología no isomórfica). Lo que encontramos es mucho más que una yuxtaposición de elementos lado a lado. Es un vínculo *activo*, en que los efectos de la limitación y la modificación pueden circular efectivamente, modificando ambos patrones en una complementariedad fructífera. Creo que este análisis de la trenza triple de la conciencia temporal ha introducido una novedad substancial, que

renueva el análisis fenomenológico más allá de sus tradicionales instrumentos descriptivos y, al mismo tiempo, une explícitamente el campo neurobiológico con el fenomenológico.

En primer lugar, esto nos permite resolver la paradoja cognitiva, puesto que un sistema dinámico no tiene "representación" de su pasado en su punto actual en el tiempo. Pero el pasado actúa en el presente en una causalidad que es mejor descrita como "invisible" o "leve", puesto que sólo puede ser revelada por reconstrucciones geométricas a partir de información bruta. En este sentido, la retención del nombre introducida por Husserl, así como el "desplazamiento", son términos descriptivos adecuados que aquí adquieren un sentido más completo. El estado presente no sería lo que es, excepto por su pasado, pero el pasado no está realmente presente (es una causalidad no lineal) y no está representado, por lo que no hay una paradoja cognitiva.

En segundo lugar, también hemos visto que la explicación presente resuelve la paradoja constitucional de la temporalidad, una dimensión de temporalidad que ha recibido mucha menos atención. Esta es la disposición para la acción que aparece en la temporalidad en la dimensión afectiva. Esta dimensión es nueva porque no sólo constituye la presencia-retencional, sino también las disposiciones sedimentadas de un ser vivo. Así, estas disposiciones van en una temporalidad mayor de su historia y situación, éstas son pre-noéticas, para utilizar de nuevo la adecuada descripción entregada recientemente por Gallagher (1998). De esta manera, el modelo dinámico no sólo es relevante para el tiempo estático retencional, sino también para la segmentación en estratos del pasado. Aquí, una vez más, el vínculo dinámico entre los datos fenoménicos y empíricos son cruciales, puesto que ofrecen las dimensiones complementarias del parámetro de orden que determina las trayectorias, así como el efecto de enlace de la acción que modifica el paisaje dinámico, de ahí que no exista paradoja constitucional. Si es posible que el modelo dinámico entregue un puente de naturalización a las relaciones internas entre las fases temporales de los actos mentales, las relaciones constitucionales son, de hecho, más ricas que lo que puede ser captado por la idea de retención de Husserl.

Es claro que hay mucho por hacer para transformar los instrumentos dinámicos en conceptos precisos para la neurociencia cognitiva, y hacer más precisas las trenzas triples para la temporalidad y otras dimensiones fenomenológicas importantes (tales como el espacio y la intersubjetividad). Pero al menos se puede decir que el programa de investigación neurofenomenológica emerge de este estudio, más allá de una declaración optimista, como un camino abierto para el análisis.

III

EPISTEMOLOGÍA

EL CÍRCULO CREATIVO

Esbozo histórico-natural de la reflexividad

Este es quizá el texto más evocador y literario de los ensayos de este libro, tal como fue propuesto cuando recibí la invitación de escribirlo. Lo que intento aquí evocar es el complejo de ideas y asociaciones alrededor de la circularidad y la autorreferencia, fuente de fascinación artística y matemática desde la Antigüedad, pero también concepto clave del pensamiento sobre la vida y el conocer, de este libro.

Cuando pensamos que ha abandonado definitivamente el plano de origen, recae nuevamente en él y dibuja su propio relieve en el blanco papel. Un círculo se cierra y al mismo tiempo dos planos coinciden, se superponen, se confunden. En esta coincidencia se observa que lo que deseábamos mantener en planos separados es inseparable. Nuestro sentido de orientación y nuestros sentimientos hacia aquello que forma la base empiezan a tambalearse y tenemos la impresión de encontrarnos ante una paradoja.

Este tipo de fenómenos circulares solían llamarse círculos del diablo (círculos viciosos) y eran la encarnación de aquello que debía evitarse. Por mi parte, recomendaría denominarlos "*circuli virtuosi*" o círculos creativos. Su notoria curiosidad oculta una clave para la comprensión de sistemas naturales y sus fenómenos

---

(Varela, F. (1981), Der kreative Zirkel: Skizzen zur Naturgeschichte der Rückbezüglichkeit, in: P.Watzlavick (Ed.), *Die Erfundene Wirklichkeit: Beiträge zum Konstruktivismus*, Piper Verlag, München, pp.294-309.)

cognitivos, como así también el rico mundo de sus formas. Me propongo aquí bosquejar este mundo de los círculos peculiares desde tres puntos de vista fundamentales: 1) el empírico; 2) el estructural; 3) el de la teoría del conocimiento.

1. La perspectiva empírica

En el dibujo de Escher observamos que ambas manos se dibujan mutuamente. Esto implica que establecen recíprocamente sus condiciones de creación. Más precisamente, su determinación recíproca las extrae del resto del dibujo, permitiéndoles conformar una "unidad". Dicho de otra manera, su operación (su mutuo dibujarse) establece las condiciones bajo las cuales pueden ser diferenciadas al tiempo que las destaca sobre un fondo.

El hecho de que una unidad se destaque de un fondo, como consecuencia de la operación, es una experiencia cotidiana que podemos asociar normalmente con los seres vivientes. Desde la Antigüedad se ha convenido en dar esta a esta experiencia el nombre de "autonomía". Cuando observo a un perro que camina por la calle, que cambia repentinamente de rumbo y se dirige hacia mí, es muy común que adjudique al perro la intención de saludarme. Determinar si esta adjudicación de un proceso mental está justificada o no es menos importante para mí que el hecho de que resulta *tentador* hacerlo en base al comportamiento del perro. Dicho de otra manera, el comportamiento del perro es muy difícil de explicar a menos que suponga que el perro no reacciona ante el mundo que lo rodea, es decir, como si recibiera instrucciones que apuntan a determinado resultados, sino más bien como si se tratara de perturbaciones que el perro interpreta de acuerdo con sus propios mecanismos de regulación y balance. Esta es, nuevamente, la cualidad tan peculiar a la que denominamos *autonomía*. Si mi automóvil no arranca mañana, efectivamente estaría tentado de decir que está enojado conmigo, pero dado que soy un individuo instruido, sé que una subordinación de este tipo no es posible, ya que nosotros mismos construimos la máquina.

Precisamente aquí comienzan las dificultades: Nosotros no creamos el perro y tampoco pareciera existir para un fin específico sobre el cual pudiéramos ponernos de acuerdo. La clara contraposición entre sistemas vivos, en los cuales se advierte la idea de autonomía, y los muchos otros sistemas naturales y artefactos creados por el hombre, fueron fascinantes para la biología, desde Aristóteles hasta entrado el siglo XIX, y en un grado solamente comparable a la atracción ejercida por la diversidad de los seres vivos[1].

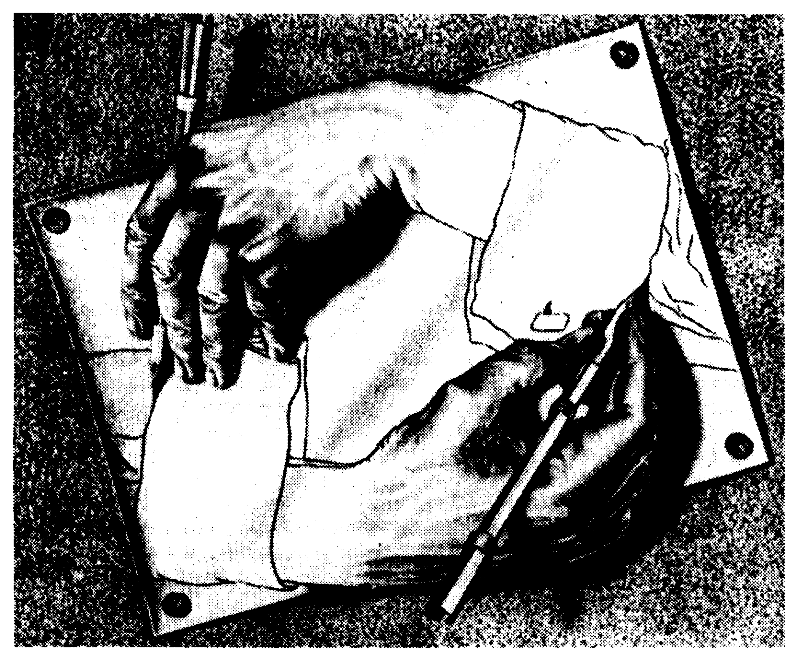


Figura 1 (M. C. Escher, *Dibujando*)

Resulta interesante comprobar que el tema de la autonomía desapareció gradualmente del discurso científico al comenzar el desarrollo de la genética y la biología molecular, en los albores del

[1] Véase, por ejemplo, Schiller J.: *La Notion d'Organisation dans l'Histoire de la Biologie*, Maloine, París, 1978.

siglo XX. Paralela y rápidamente, la técnica y la mecánica hicieron rápidos progresos y se orientaron hacia la cibernética y la teoría del control. Este es el motivo por el cual hoy día no pensamos en autonomía en el área de los sistemas naturales, sino que simplemente lo pasamos por alto sin considerar que se pueda hablar de autonomía de manera precisa. La contrapartida de la autonomía, el control, puede, en cambio, precisarse sin inconvenientes.

Naturalmente que no existe nada más misterioso en la autonomía que lo que pueda haber en el control. Lo decisivo radica en considerar a la autonomía como la expresión de un tipo de *proceso* que aparece por doquier en la naturaleza y en numerosas formas concretas[2]. Este tipo de proceso es exactamente lo que Escher presentó. (Las partes se especifican mutuamente y se fijan entre sí).

La vida se caracteriza por esta forma de vinculación con el ámbito molecular, adquiriendo su cualidad de autónoma. De esta "sopa" de moléculas se destaca una célula, porque define y fija fronteras que la separarán de todo aquello que no es ella. Sin embargo, esta determinación de fronteras se cumple mediante producciones moleculares que, por su parte, sólo fueron posibilitadas por esas fronteras. Las transformaciones químicas y los límites físicos se condicionan mutuamente: la célula se destaca de un entorno homogéneo. Si este proceso de autocreación se interrumpe, la unidad celular deja de formar una unidad y se deshace hasta formar poco a poco la sopa molecular[3].

La esencia de la organización molecular se puede representar de la siguiente manera:

---

2 Para una presentación detallada de esta idea, véase Varela F.: *Principles of Biological Autonomy*, North Holland, Nueva York, 1979.

3 Maturana H. y Varela F.: *Autopoiesis and Cognition*. Boston Stud. Phil. Sci., Vol.42. D. Reidel, Boston, 1980.

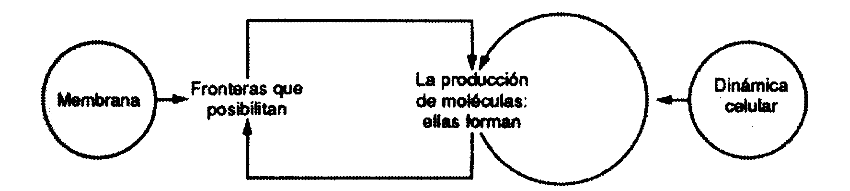


Figura 2

Esta configuración es determinante: las operaciones conforman un *círculo cerrado*, motivo por el cual los productos se encuentran en el mismo plano que los proceso de producción. Dentro de esta organización pierden sentido las diferenciaciones usuales entre productor y producto, entre comienzo y final o entre *input* y *output*. Poco podemos decir sobre el origen de las células, pero los resultados de las últimas investigaciones concuerdan sustancialmente con la idea de que el cierre de un circuito generacional como el que acabamos de bosquejar resulta de una condición indispensable[4]. Una vez que existen tales unidades autónomas surge todo un nuevo dominio: la vida como la conocemos hoy. Este tema fundamental de la interacción de los circuitos de procesos de producción molecular admite numerosas variaciones, así como también muchas materializaciones específicas que dan como resultado una infinidad de células distintas.

Es posible que las células modernas sean el fruto de una simbiosis de unidades, originariamente autónomas. Tal es el caso de las mitocondrias, cloroplastos y otros organelos celulares, los cuales hoy en día conservan vestigios de su autonomía original[5]. Incluso en nuestros días, encontramos algas y hongos que conforman sus cambiantes recíprocamente y suministrándose alimentos. Por consiguiente, las células pueden interactuar formando nuevas células autónomas. Todos los organismos multicelulares fueron creados en similares circunstancias.

---

[4] Eigen, M. y Schuster, P.: *The Hypercircle*. Springer, Berlin, 1979.

[5] Margulis, L.: *The Evolution of Eucaryotic Cells*. Freeman, San Francisco, 1980.

El fenómeno básico, en todos estos casos, es el mismo: Elementos de diversos planos se reúnen operacionalmente y forman una unidad a raíz de su interacción circular. Si este proceso se interrumpe, se destruye esta unidad. La autonomía nace en esta intersección. El surgimiento de la vida no es un mal ejemplo para esta ley general.

2. La perspectiva estructural

"Esta frase es falsa si se la añade a sí misma entre comillas", es falsa cuando se la añade a sí misma. Este *Koan* de Quine[6] es una expresión jocosa de una dificultad a la cual se enfrentan hace largo tiempo la lingüística y la matemática. A partir de que el cretense Epiménides se le ocurrió decir aquello de que "todos los cretenses son mentirosos", la singular cualidad humana de la reflexividad (autorreferencia) fue la causa de permanentes dolores de cabeza[7]. Esta singular cualidad reposa sobre un postulado que determina que las afirmaciones sobre algo no deben ser elemento constitutivo de ese algo.

Afirmaciones como la de Epiménides y Quine lesionan obviamente este postulado.

En todos los casos similares en los que existen confusiones lingüísticas, es evidente cierto parecido familiar con el grabado de Escher, así como la formación de las células y la autonomía. En todos los casos, se trata de movimientos con los cuales aquello que debería quedar separado se entrecruza (en el caso de Quine y de Epiménides serían los planos de significación), de manera que dos planos se confunden en uno solo y, a pesar de todo, siguen siendo diferenciables.

---

[6] Quine, W. O.: *The Ways of Paradox and other Essays.* Harvard U. Press, 1971.

[7] Para un debate completo, véase Hughes, Patrick y Brecht: *Visions,* Doubleday, Nueva York, 1975, Vieweg, Braunschweig, 1978, y especialmente el libro de Hofstadter, D.: *Gödel, Escher, Bach,* Basic Books, Nueva York, 1979.

Sin embargo, resulta interesante comprobar que aquello, que en el ámbito molecular aparece como complejo pero entendible, adquiere en el ámbito lingüístico la significación más profunda de una *paradoja.* Es más difícil saltar fuera de la necesidad de permanecer en un determinado plano de significación y considerar sencillamente toda la oración como una unidad. Una paradoja es exactamente esto, lo que permanece incomprensible si no lo examinamos saliendo de ambos planos mezclados en la estructura de la paradoja. Quine y Epiménides siguen siendo paradójicos en la medida en que no estoy dispuesto a abandonar la necesidad de elegir entre lo verdadero y lo falso, así como a reconocer en la reflexividad de la premisa una forma determinada de fijación de su significado. Esto implica que la frase existe dentro de un ámbito más amplio y sólo se torna paradójica en la medida en que se la proyecte sobre un ámbito más restringido, en el cual debe ser obligatoriamente verdadera o falsa.

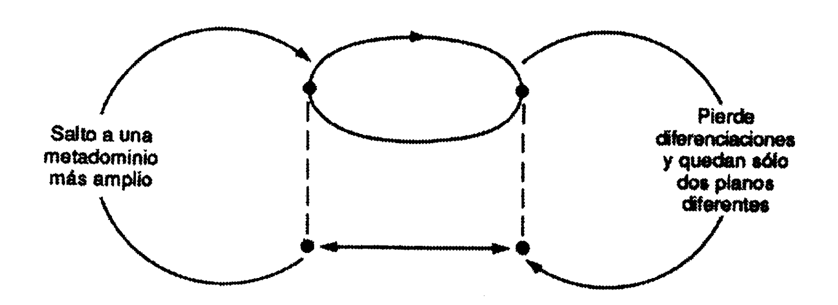


Figura 3

Este es, supongo, el motivo por el cual aparece la paradoja en situaciones como la de los ejercicios zen, en los que justamente debería aprenderse a saltar a un plano cognitivo superior para poder observar en este nuevo plano sus pensamientos y conceptos valorativos en forma imparcial. Mientras el aprendiz se mantenga atado a uno u otro plano, a una predilección o juicio, a lo bueno o lo malo, a lo positivo o lo negativo, a lo espiritual o lo mundano, la meta de la enseñanza no se

habrá alcanzado. Un buen maestro es aquel que puede transmitir vívidamente la reflexividad y lo entreverado de la situación hasta que el aprendiz se vea obligado a extraerse de ella.

Quizá la prueba más interesante y famosa de la fecundidad en el ámbito de la reflexividad en el lenguaje y en la matemática sea el teorema de Gödel, ejemplo del cual quisiera extraer conclusiones en relación a la circularidad y cierre.

La intuición de Gödel está (también) representada perfectamente en el grabado de Escher. Gödel (y sus contemporáneos) querían determinar si los lenguajes formales podían autoanalizarse; dicho de otra manera, querían ver si se los podía analizar exclusivamente con sus propios medios. Para ello debemos ocuparnos por lo menos de aquellos lenguajes matemáticos que contengan los números naturales y que puedan hacer afirmaciones sobre los números. Claro está que los números no son afirmaciones matemáticas, sino objetos matemáticos a los cuales uno puede referirse en un lenguaje matemático apropiado para este fin. La ocurrencia genial de Gödel consistió en entrecruzar el plano del lenguaje sobre los números con los números mismos. ¡Realmente un círculo peculiar! Para ello adjudicó un signo lingüístico a cada número, de manera tal que también a las series de números (es decir, afirmaciones sobre números) correspondía una cifra. Los detalles no nos interesan aquí[8], pero el punto central del lenguaje construido de esta manera por Gödel es el siguiente:

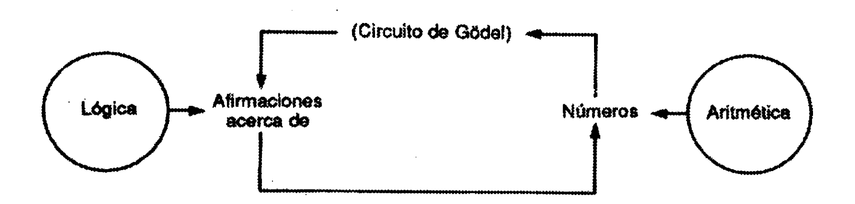


Figura 4

[8] Para mayores detalles, véase Nagel, E. y Newman J. R.: *Gödel's Roof*, Nueva York, University Press, 1965, así como Hofstadter, *obra citada*.

Cuando, en este sentido, se entrecruzan claramente diversas áreas, no resulta dificultoso conformar premisas reflexivas como la de Quine. Gödel manifiesta este mismo pensamiento, diciendo: "Esta afirmación es indemostrable"(ni su veracidad ni su falsedad pueden demostrarse). La sola existencia de tal afirmación demuestra que todos los sistemas formales, de hecho suficientemente ricos como para contener números y la aritmética, contienen afirmaciones llenas de sentido y perfectamente definidas de las cuales no se puede inferir si son verdaderas o falsas. Por eso se dice de estos sistemas formales que son *incompletos*. El hecho de que existan en afirmaciones indecibles, justamente dentro del ámbito central de la matemática, produjo gran insatisfacción entre los matemáticos. Desde nuestro punto de vista, en cambio, los resultados a los que arriba llega Gödel se interpretan de manera muy diferente. No como una prueba de la limitación, sino como un caso más que demuestra que la circularidad puede llevar a la constitución de un dominio autónomo, formándose una unidad que especifica un dominio abarcador y más amplio. En el caso de Gödel, apenas se completa su circuito y los planos se cruzan, aparece una unidad en el universo lingüístico. La comparación con el ejemplo biológico es obvia.

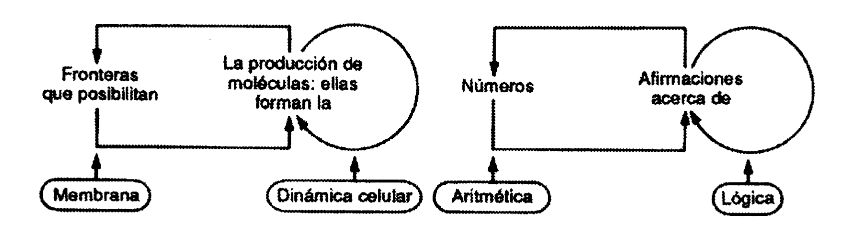
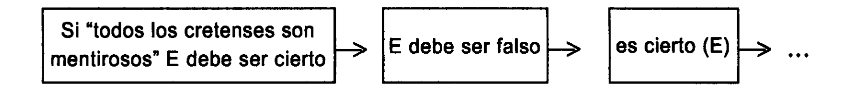


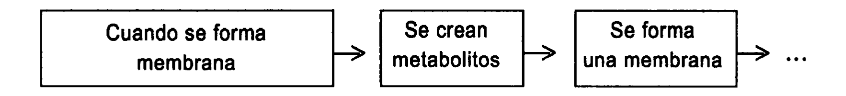
Figura 5

Observen un momento qué es lo que se desarrolla *dentro* de estos curiosos circuitos. En el caso de Epiménides, la afirmación es –cuando aceptamos que es verdadera–, falsa. Si es falsa, consecuentemente debe ser verdadera. En su estructura hay una

oscilación entre lo que antes se encontraba separado. Lo podemos graficar de la siguiente manera:



En lo que a la célula se refiere, podemos abrir el círculo y obtenemos igualmente una estructura que se prolonga al infinito.



En la unidad de un circuito operacional, aquello que aparece como una *conducta coherente* y diferenciable (ya sea en el área de la significación o en el de las estructuras moleculares) es, en realidad, peculiar. Por una parte, este comportamiento actúa como una característica de la unidad. Por otra parte, vemos que en el intento de determinar el origen de esta característica por medio de sus cualidades encontramos solamente una reiteración sin fin de lo que es siempre igual, que no comienza en ninguna parte y que no termina en ninguna parte. La coherencia está *distribuida* en un círculo que se reitera constantemente, que repite indefinidamente, pero que es finito como circuito, dado que podemos observar sus efectos o sus resultados como característica de una unidad.

Quisiera ilustrar el mismo pensamiento más virtualmente. Consideremos un triángulo. Dividimos cada lado en tres partes, las cuales unimos luego de tal manera que formen una estrella de seis puntas. Procedemos ahora a dividir de igual manera cada lado de la estrella. Este proceso se repetirá con cada nuevo lado creado, *ad infinitum*. La figura así creada se asemeja a un cristal de nieve y es

inmediatamente comprensible pues posee una forma coherente. Pero lo que percibimos es como un antepasado mítico que nunca fue totalmente dibujado y que sólo puede ser intuido como tendencia de una repetición ininterrumpida. Resulta interesante el hecho que figuras como éstas tienen, en base a su construcción geométrica autorreferencial, dimensiones que no son las tradicionales. En el caso del triángulo de más arriba, la dimensión es superior a 1 pero inferior a 2, exactamente 1,2618. Dado que la dimensión es un número (fraccionario) a tales figuras se las denomina *fractales*[9].

3. La perspectiva cognitiva

Hemos presentado dos perspectivas paralelas, en las cuales la formación de círculos operacionales, mediante el acto aparentemente inofensivo de la reflexividad, crea un ámbito completamente nuevo, así se trate de células y de seres vivos o de lenguajes e indeterminación.

Ahora debemos dar el próximo paso en nuestra investigación de la historia natural de la reflexividad y analizar el otro caso fundamental, cuyo cierre hace variar totalmente el cuadro: descripciones de *nosotros mismos*, nuestro *propio* conocer.

Precisamente, en la observación de nuestros propios conocimientos, reunimos los tópicos principales de las dos perspectivas anteriormente mencionadas.

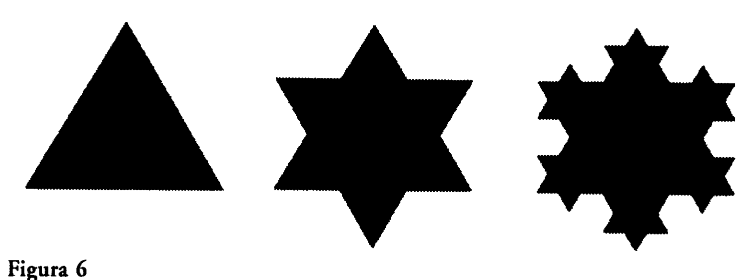


Figura 6

[9] Eigen, M. Y Schuster, P.: *The Hypercircle*. Springer, Berlín, 1979.

Por una parte, nuestra cognición ocurre en el sustrato biológico de nuestro cuerpo. Por otra parte, nuestras descripciones son capaces de autodescripciones. Gracias a nuestro sistema nervioso, superponen ambos modos de cierre y forman así aquella vivencia que es la más familiar y, al mismo tiempo, la más inasible: nosotros mismos.

Está claro que el sistema nervioso es una parte integrante de nuestra unidad como seres biológicos, como unidades autónomas que somos. Lo que no es tan evidente es que el propio sistema nervioso es autorreflexivo de varias maneras básicas[10].

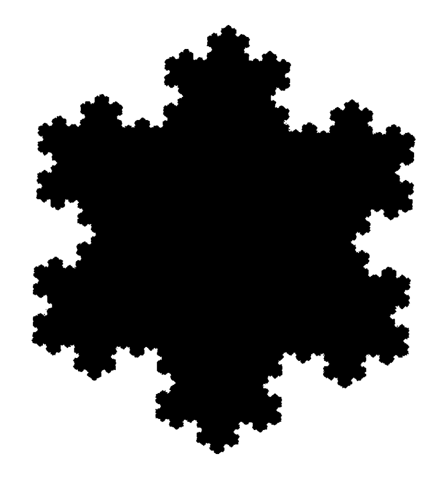


Figura 7

Esto es así sobre todo porque no hay efecto del sistema nervioso (capacidad de movimiento, secreciones internas) que no tenga un efecto directo sobre una superficie sensorial. De la misma manera que una neurona actúa sobre otra por estrecha vecindad de sus superficies a través de una sinapsis, un grupo de músculos actúa sobre el sistema sensorial del cuerpo por efecto reflexivo sobre una sinapsis sensorial y motora.

---

[10] Maturana, H.: *Biologie der Kognition.* Paderborn, 1975.

Un movimiento reflejo de la rodilla es provocado por la tracción de un tendón, que ejerce una tracción sobre propiorreceptores y provoca una modificación de la actividad de las neuronas motoras en la médula espinal, todo lo cual conduce a una contracción muscular en dirección opuesta a la tensión del tendón. Los efectos motores tienen consecuencias sensoriales y los efectos sensoriales tienen consecuencias motoras. Este principio de referencia tiene validez universal:

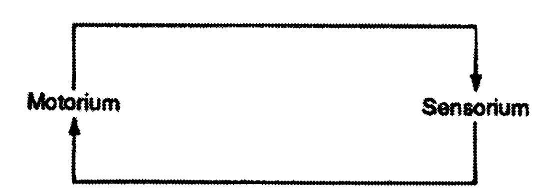


Figura 8

Pero el sistema nervioso, en un sentido más sustancial, es un circuito cerrado. Tan pronto se traspone el umbral de lo sensorial o de lo motor, los efectos que éstos ejercen no tienen una dirección única, como en una calle de una sola vía. Se parecen más bien a la aparición de otro comprador en la sala de la Bolsa de valores. Si, por ejemplo, siguiéramos la excitación creada en la retina en su trayecto hacia la zona de la corteza cerebral (lóbulo occipital), podríamos comprobar que en cada fibra que llega a la corteza, proveniente de la retina, desembocan en este punto 100 fibras provenientes de otros lugares del cerebro[11]. La actividad de la retina suministra así sólo una modulación de aquello que se desarrolla en el sistema nervioso central.

Pero esto no es todo. Si bien los estímulos eléctricos se propagan sólo en una dirección, muchos otros estímulos químicos del eje neuronal se desplazan en sentido contrario, de manera que

---

[11] Braitenberg, V: *Gehirngespinste: Neuroanatomic für Kybernetisch Interessierte.* Springer, Berlín, 1973.

las vías de transmisión son siempre "calles de doble tránsito" en el sistema nervioso. Así, por ejemplo, puede incorporarse en el extremo del axón neuronal una sustancia reguladora que viaja hacia el cuerpo celular y actuar por medio de un impulso eléctrico sobre la neurona procedente a través de una sinapsis. En el sistema nervioso existen múltiples efectos recíprocos de este tipo, cuyo funcionamiento sólo ahora se está empezando a estudiar[12].

En el siguiente esquema se observa la descripción descrita:

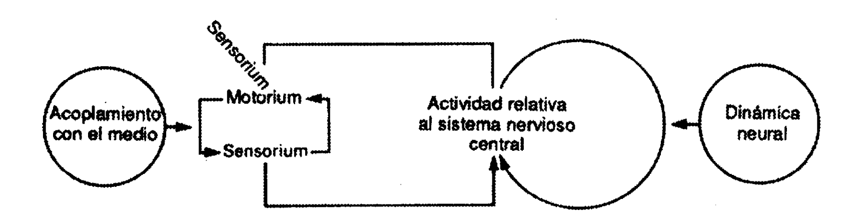


Figura 9

En esta visión del sistema nervioso observamos una conducta determinada cuando, por haberse cerrado el circuito de esta estructura total de conexiones recíprocas, se logra coherencia. Se podría analizar esta coherencia dividiendo el proceso en sus partes constitutivas. Podríamos, por ejemplo, comenzar con la visión que tiene su origen en el ojo y seguir luego todos los trayectos que llevan desde el ojo hacia la corteza cerebral y luego desde la corteza al tálamo y al lóbulo anterior, etc. Finalmente, habríamos descrito un círculo completo y, de hecho, podríamos seguir girando indefinidamente en él. La conducta se refiere, como en el ejemplo del *fractal* –antecesor mítico de este proceso infinitamente recurrente– a sí misma.

Si tomamos en serio lo que dijimos acerca del sistema nervioso, debemos reconocer que nuestra experiencia personal se

[12] Para un nuevo enfoque, véase Dismukes R. K.: The Brain Behaviour. *Science 2,* 1979, pág. 409.

origina de la misma manera. De esto se extraen dos conclusiones de enorme importancia:

Primero: No podemos salir del mundo determinado por nuestro cuerpo y nuestro sistema nervioso. No existe otro mundo excepto el que experimentamos por medio de estos procesos, que son premisas para nosotros y hacen de nosotros lo que somos. Nos encontramos dentro de un dominio cognitivo del cual no podemos salir o decidir dónde comienza o cómo se crea.

Segundo: (y no menos importante): No podemos retrotraer una experiencia dada de una manera única e irrepetible a sus orígenes. Cada vez que intentamos rastrear los orígenes de una percepción o de una idea, chocamos contra un *fractal* que retrocede permanentemente ante nosotros. No importa dónde investiguemos, tropezamos siempre con la misma multiplicidad de detalles y vinculaciones recíprocas. En todos los casos se trata de la percepción de la percepción, etc., o de la descripción de la descripción, etc. En ningún lugar podemos arrojar el ancla y decir: "De aquí partió esta percepción y de esta manera se desarrolló". En nuestra percepción del mundo olvidamos todo aquello que aportamos para percibirlo de este modo precisamente porque estamos incluidos a través de nuestros cuerpos en el proceso circular de nuestros comportamientos. Al igual que el joven del dibujo de Escher "Galería de cuadros", vemos un mundo que se convierte en el sustrato que nos produce, cerrando así el círculo y haciendo que se entrecrucen las áreas. Al igual que en el grabado de Escher, no hay salida hacia ninguna parte. Si intentáramos salir, nos encontraríamos nuevamente en un círculo sin fin que se perdería en un punto central, dentro de un espacio vacío.

Conforme a la tradición[13], la experiencia es objetiva o subjetiva. El mundo existe y nosotros lo podemos ver tal como es (objetivamente) o bien lo vemos a través de nuestra subjetividad. Si

---

[13] La diversidad de tradiciones debería ser determinada con mayor detenimiento. En este sentido resulta muy significativa la fenomenología con sus diferentes ramificaciones. Pero aquí se habla, sin embargo, del sentido común dominante.

seguimos el hilo conductor de la reflexividad y de su historia de la naturaleza, podemos ver esta intrincada pregunta desde otro punto de vista: el de la *participación* y de la *interpretación,* en el cual el sujeto y el objeto están inseparablemente unidos entre sí. Esta interdependencia se pone en evidencia por el hecho que no puede comenzar en ninguna parte con una representación pura y no contaminada de lo uno o de lo otro y cualquiera sea el lugar por el que resuelva comenzar se las tendrá que ver hasta cierto punto con un *fractal* que reproduce exactamente lo que él hace, es decir, describirlo. De acuerdo con esta lógica, nuestro comportamiento en relación al mundo es igual al que tenemos ante un espejo, el que ni nos podrá decir cómo es el mundo ni cómo es. El nos muestra que es *posible* que seamos como somos y que actuemos como hemos actuado. Nos muestra que nuestra experiencia ha sido *viable.*



Figura 10 (M. C. Escher, *Galería de Cuadros*)

Es fascinante que el mundo sea así de plástico, ni subjetivo ni objetivo, ni unitario ni separable, ni dual ni inseparable. Esto apunta a la *naturaleza* del proceso, que podemos percibir en la totalidad de su calidad formal y material, así como también a los *límites* fundamentales de aquello que podemos comprender de nosotros mismos y del mundo. Demuestra que la realidad no está constituida sencillamente a nuestro antojo, porque esto significaría suponer que podemos elegir un punto de salida desde adentro. Prueba además que la realidad no puede entenderse como algo objetivamente dado, como algo que recogemos, porque esto significaría suponer un punto de partida externo. Demuestra de hecho una ausencia de fundamento sólido de nuestras experiencias, en las cuales nos son suministradas determinadas regularidades e interpretaciones, fruto de nuestra historia conjunta como seres biológicos y sociales. Dentro de estas áreas de historia común que reposan sobre acuerdos tácitos, vivimos en una aparentemente interminable metamorfosis de interpretaciones que se suceden entre sí[14].

Se nos revela un mundo en el cual lo in-fundado puede convertirse en base de comprensión de que el antiquísimo ideal de la objetividad y de la comunicación, entendidos como la progresiva eliminación de los errores, en beneficio del aumento de la coincidencia, medido en sus propias escalas científicas, es una quimera. Haríamos mejor en aceptar de lleno la situación rotundamente distinta y difícil que nos plantea, a saber de que vivimos en un mundo en el que nadie puede pretender comprender las cosas, en sentido amplio, de mejor manera que los demás. Lo notable es que el mundo empírico de los vivientes y la lógica de la autorreferencia, así como las enseñanzas de toda la historia natural de la reflexividad, nos enseñan que la ética, la tolerancia y el pluralismo nos liberan de nuestros propios valores y percepciones, para respetar las percepciones y los valores de los demás. Esto constituye, en definitiva, el conocimiento y, al mismo tiempo, su punto final. En este punto, los hechos son más claros que las palabras.

[14] La expresión filosófica más concisa que he encontrado para esta conclusión es la filosofía Madhyamika de la escuela medieval hindú. Véase, por ejemplo, las útiles explicaciones de Streng, F.: *Emptiness: A study in religious meaning*, Abdington Press, Nueva York, 1967.

COMPRENDIENDO LOS ORÍGENES
Una introducción

El origen aparece en plural porque hay una continuidad estructural notable a través de múltiples dominios desde la vida, el conocer, el orden social y la moneda. Se trata de un origen sin substancia, circularmente originado como proyección exógena y que se asocia a las grandes discusiones de la filosofía posmoderna. Este texto fue escrito como introducción a un simposio enteramente dedicado al tema.

Quisiéramos comenzar con la siguiente observación: las humanidades y las ciencias "duras" (término aquí referido específicamente a la biología y a buena parte de las ciencias cognitivas) difieren considerablemente en sus ambiciones concernientes a las "grandes preguntas". Las ciencias duras se han vuelto más audaces que nunca en sus proposiciones acerca de cómo se formó el cosmos y se originó la vida, y de qué manera evolucionaron las especies. Por el contrario, para las humanidades ha sido un tiempo de dispersión, de fragmentación, de una diseminación que resiste cualquier intento de integración a gran escala. El tiempo de las "grandes teorías" parece haber quedado bastante atrás.

Esto no siempre fue así y podemos señalar 1939, año de la muerte de Freud, como el punto de quiebre. Hasta entonces, la tradición de la antropología religiosa no dudaba en postular un

---

(Dupuy,J.P. and F.Varela (1991), Creative circularities: Understanding origins, in: F.Varela and J.P.Dupuy (Eds.) *Understanding Origin*, Stanford University Intl.Symposium, Boston Studies Phil.Sci., Kluwer, Boston, pp. 1-27)

origen a las culturas humanas, una cuestión que era inseparable de aquella del origen de lo religioso. Las figuras centrales en esta tradición —Fustel de Coulanges, Robertson Smith, Frazer, Durkheim, Hocart y Freud— compartían la hipótesis de que cada institución humana está fundada en lo religioso, o para ser más precisos, en lo ritual. Una multiplicidad de factores puede invocarse para explicar por qué razones esta tradición desapareció rápidamente después de la Segunda Guerra Mundial, y por qué las escuelas posteriores la consideraron como simplemente irrelevante. Podemos pensar, por ejemplo, en el movimiento estructuralista francés y su postulado de que las estructuras 'están siempre previamente ahí'. O la postura más positivista de la antropología anglosajona, representada por ejemplo por Evans-Pritchard. Pero, sobre todo, queremos reflexionar sobre el impacto que tuvieron sobre las humanidades los trabajos de Nietzsche y Heidegger, seguido por la escuela de Jacques Derrida cuyo centro es la así llamada "desconstrucción de la metafísica occidental".

La principal interrogante de la metafísica, según Heidegger, es aquella referida al origen de las entidades: ¿por qué hay algo en lugar de nada? Enfrentada a esta pregunta, la metafísica occidental responde con el principio postulado por Leibnitz de la razón suficiente: cada efecto tiene una causa, ésta surge de otra causa y así sucesivamente hasta llegar a la causa primera, que es autosuficiente, completa y la causa de sí misma, esto es, Dios. Este es el argumento onto-teológico: a la pregunta ¿por qué? uno responde con el postulado de una entidad fundamental, lo cual establece la diferencia ontológica entre el Ser y las entidades. De este modo, el misterio del "hay" desaparece bajo el peso de la respuesta.

Heidegger, por el contrario, no propone ninguna respuesta, al partir de la desconstrucción de la pregunta misma: no es necesario buscar una causa sino permitir que el misterio se despliegue totalmente; la respuesta es un signo de interrogación. "La rose est sans pourquoi" (Rimbaud); es decir, no tiene razón ni causa. La desconstrucción derridiana sigue a Heidegger de cerca. Se concentrará en aquellas creaciones filosóficas como Naturaleza, Lenguaje, Razón, Origen, Significado, Verdad y Sujeto, que aparecen como completas, autosuficientes y la causa de sí mismas. El punto es

desconstruir la pretensión del Logos de afirmarse a sí mismo como completo y autosuficiente, la ambición de la filosofía de tener acceso inmediato a la verdad pura (*aletheia*), la ilusión de dominio por parte del sujeto humano que se coloca a sí mismo en el lugar de Dios. En su "desconstrucción de la metafísica occidental", la empresa intelectual iniciada por Derrida y sus numerosos seguidores en la senda de Heidegger, sistemáticamente desprestigia el Concepto que, al igual que el burgués autorreferente ridiculizado por Marx en L*a Sagrada Familia* "se hincha hasta el punto de considerarse a sí mismo como un átomo, es decir, como un ser incapaz de ninguna relación, suficiente en sí mismo, carente de necesidades, absolutamente completo, en un estado de total felicidad".

La principal herramienta usada por la desconstrucción es lo que Derrida llama la *lógica del suplemento*. Como dice René Girard en el texto que sigue, esta lógica "refleja [la] inaptitud general del ser humano para el auto-centrismo, [la] falla del narcisismo individual y colectivo, y el resultante temor de y fascinación por la otredad". Esta lógica es aquella a través de la cual cada texto filosófico se desconstruye a sí mismo. Cada vez que un término aparece en un texto teórico que señala a un Logos, a un Concepto, como autosuficiente, se genera un círculo vicioso que debilita esta pretensión de autonomía desde adentro. Esto ocurre porque otro término, que se supone es secundario y subordinado y que no debería ser otra cosa que una derivación o complicación del Concepto primario (por ejemplo: cultura, escritura, forma, etc.), aparece como indispensable para la constitución de éste último. El origen aparece como completo y puro, pero, si careciera del suplemento, que sin duda se deriva del mismo, perdería toda consistencia. De este modo, el término secundario aparece al mismo tiempo como perfectamente dispensable y perfectamente indispensable. Incluso la totalidad aparentemente más perfecta sufre ineludiblemente de una carencia constitutiva.

La lógica del suplemento puede entonces mostrarse en la forma de una causalidad circular que unifica a dos términos a pesar del hecho de que uno alega ser jerárquicamente superior al otro, como se indica en la Figura 1.

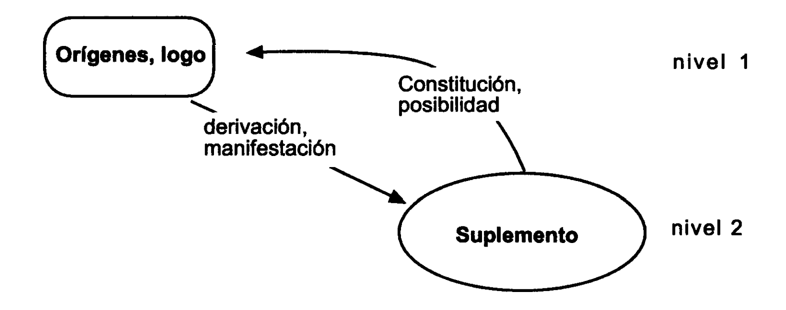


Figura 1

Es preciso señalar que la desconstrucción de una oposición jerárquica no es lo mismo que su simple remoción. La dimensión jerárquica debe estar presente de una u otra manera. Tampoco consiste la desconstrucción en simplemente invertir la oposición jerárquica, en una permuta de sus términos superior e inferior. Considérese el ejemplo, especialmente importante para Derrida, de la oposición jerárquica entre la filosofía y la escritura. ¡La filosofía devalúa la escritura precisamente porque está escrita! La escritura constituye una amenaza a la filosofía de la misma manera que el dinero lo es para la economía, porque es un obstáculo, una barrera en el camino de acceso al significado y al valor. Ya que el ideal de la filosofía es alcanzar la verdad sin mediación, debe entonces negar el único medio que tiene para expresarse: la escritura. Puesto de manera burda, la filosofía escribe W:

W: "*Esto no es escritura*"

forma obvia de una paradoja autorreferente. Volveremos a este tema de la autorreferencia frecuentemente.

Existe muy poco intercambio entre la actividad de la desconstrucción y la producción científica. Los desconstructores pueden saborear la ilusión de tener en sus manos una no-lógica, ya que no solamente parece eludir a primera vista la formalización

científica, sino que además debilita cualquier postulado acerca de un origen. Simplemente tómese cualquier texto científico referido al origen de cualquier dominio —vida, significado, lenguaje, dinero— y uno podría apostar a que este discurso se desconstruiría de acuerdo con la lógica del suplemento.

Consideremos un importante ejemplo que coincide muy bien con esta expectativa: la biología molecular y celular. Parecía ser que esta disciplina es un modelo de éxito en la reducción de la vida a una química macromolecular, principalmente a través del descubrimiento del código genético y la noción de una programación celular que se supone subyace a todo desarrollo ya que (literalmente) escribe al organismo a medida que se despliega en su ontogenia. Sin embargo, luego de una fase inicial de euforia con la idea, se ha vuelto bastante claro —y los biólogos moleculares fueron los primeros en hacer ver esto— que si uno toma la noción de programa genético en sentido literal, cae en un extraño circuito cerrado: uno tiene un programa que para ser ejecutado necesita de su propio producto. Cada paso de mantenimiento y transcripción del ADN es mediatizado por proteínas, las cuales son precisamente aquello que es codificado. Para llevar adelante el programa ¡éste debe ya haber sido ejecutado! Podemos ilustrar esta situación como se indica en la Figura 2.

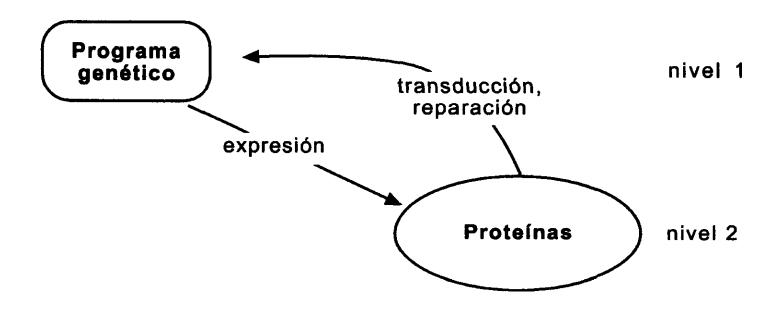


Figura 2

Ahora bien, esto no es algo que preocupe excesivamente a un biólogo. Para él, la paradoja está resuelta, dado que cada célula ha sido derivada de otra célula y, por lo tanto, una ontogenia individual comienza en el huevo mismo maternal fertilizado. Desde un punto de vista más teórico, el tema se complica, ya que en algún punto uno debe llegar al origen de la vida e iniciar la cadena de individuos autónomos. Estos esfuerzos por fundamentar la autonomía de lo viviente han tomado precisamente la forma de un circuito cerrado aparentemente paradójico o *autopoiesis*[1]: la lógica de la célula es una de autogeneración por una determinación circular entre sus límites y su dinámica, que produce los límites y al mismo tiempo es posible gracias a ellos, como se indica en la Figura 3.

Para los teóricos, esto presenta un profundo interés: la observación de que bajo ciertas condiciones un conjunto de componentes puede enlazarse en una causalidad intrincada y circular —o para ser más técnicos, mediante un cierre operacional[2]:— de tal manera que existe un nuevo nivel emergente que no es ni reducible a la suma de sus componentes ni independiente de sus productos. El circuito cerrado aparentemente paradójico entre dos niveles mostrará a primera vista una relación jerárquica (por ejemplo, la célula es una unidad versus su dinámica química), pero ambos están irremediablemente entremezclados. Entonces, la forma de esta lógica es la de dos niveles que deben mantenerse independientes y, sin embargo, están innegablemente entrelazados.

Este tipo de autorreferencia activa está en el corazón de la materia que deseamos abordar en este libro. En otras palabras, uno de nuestros principales argumentos es que las cualidades de los procesos operativos a distintos niveles comparten características comunes, siendo la más importante la constitución de una entidad unitaria (ya sea un sistema de células, un idioma, o un sistema monetario) que parece exterior a sus componentes, y que sin embargo es engendrado

---

1 H. Maturana y F. Varela, *Autopoiesis and Cognition* [Boston Studies, vol. 42] D. Reidel, 1980; L. Margulis, *Simbiosis in Cell Evolution*, Freeman, 1982.

2 F. Varela, *Principles of Biological Autonomy*, North-Holland, Nueva York, 1979.

en forma endógena por el entrelazamiento de aquellos mismos componentes. En este sentido, lo que puede aparecer como exógeno (el Valor, lo Otro), puede ser perfectamente consistente con el hecho de ser constituido endógenamente si uno mira la lógica inclusiva completa de la situación.

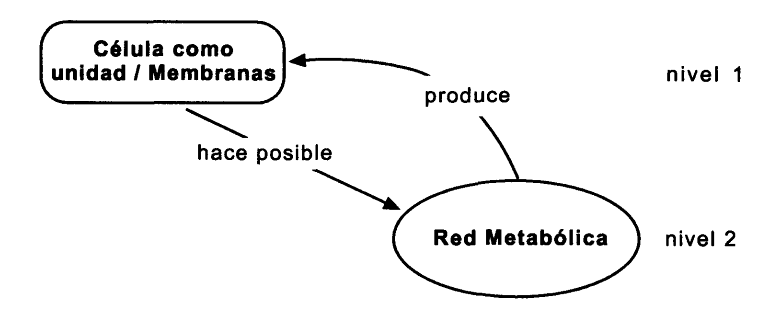


Figura 3

Ahora bien, es extraordinario el hecho de que aquellos teóricos de los sistemas naturales autónomos que han intentado abordar los fundamentos originales de la vida terminan con una lógica que es una reflexión especular de la lógica derridiana del suplemento ¡que es la principal arma para llegar precisamente a la negación de toda pretensión de un fundamento autónomo y una auto-suficiencia! Uno de los principales propósitos de la reunión y una pauta principal en este libro ha sido el tratar de derrumbar los muros académicos que han mantenido separadas estas dos modalidades independientes de trabajo, confrontarlas y tratar de buscar una mayor claridad. Nuestro enfoque en esta tarea ha sido el de tomar las nociones específicas de origen en la vida, mente y sociedad como casos de estudio. Hagamos ahora una pausa antes de retornar a los principios más generales que hemos abordado en esta sección con el fin de considerar las pautas básicas de los casos de estudio incluidos aquí.

Parecía inevitable detenerse un momento sobre el tema del origen del orden social visto desde el punto de vista privilegiado de la antropología fundamental de René Girard.

El enfoque de Girard es de hecho el único existente que se conecta retrospectivamente con la antropología religiosa clásica, desafiando los interdictos tanto del estructuralismo como de su deconstrucción. No duda en enfrentar la pregunta por el origen de lo sagrado y, a través de éste, el origen de todas las instituciones sociales y culturales. Como veremos, sin embargo, el núcleo de la lógica utilizada por Girard no es otro que el de la desconstrucción y la teoría de los sistemas autónomos.

No podemos resumir aquí el vasto y complejo trabajo de Girard, el cual se ha difundido con el correr de los años a través de varias presentaciones que constituyen desarrollos progresivos[3]. Hay dos hipótesis fundacionales que son importantes para nosotros aquí, siendo la segunda, en principio, derivación de la primera. La primera, y más importante, es la hipótesis del mimetismo: los hombres se imitan unos a otros en su deseo; nunca deseamos otra cosa que lo que el otro desea. Esta mimesis de apropiación termina inevitablemente en conflicto y violencia; lo otro muta a partir de un modelo para convertirse automáticamente en un rival, en un obstáculo.

Esto conduce a la segunda hipótesis, la de la victimización: todas las sociedades primitivas supuestamente han vivido un evento primordial y fundacional: en el clímax de la guerra de todos contra todos desatada por la mimesis de la apropiación, se habría producido una polarización de toda la violencia para enfocarla arbitrariamente en un miembro del grupo. Estas exclusiones sacrificiales, vividas en la ignorancia de dicho mecanismo, constituirían la fuente de lo sagrado, de la cultura, de todas las instituciones humanas.

---

3 Las que tenemos en mente aquí son *Deceit, Desire and the Novel,* The Johns Hopkins U. Press, 1965; *Violence and the Sacred* (bid., 1977; *Des choses cachées depuis la fondation du monde*, Grasset, 1978; *Le Bouc émissaire*, Grasset, 1982.

Ahora bien, ciertamente es posible leer la tesis de Girard a partir del ventajoso punto de vista de la desconstrucción —que no es lo mismo que desconstruir a Girard con las herramientas del desconstructivismo. Esto es lo que McKenna hace en su texto: "El escenario originario de Girard puede adaptarse a una crítica postestructuralista porque propone que origen y *différance*[4] son una misma cosa, que la representación es el subproducto de una *différance* originaria o un deseo mimético en lugar de la representación de una presencia originaria, de un origen de cualquier tipo". Uno puede argüir junto con McKenna que en *ambos* Girard y Derrida "al principio está la imitación, no el origen". Esto es verdad en cierto sentido para el mecanismo mimético: la situación original no puede ser la de un sujeto *A* que imita el deseo de otro sujeto *B* que desea en forma autónoma, ya que la mimesis es universal. Entonces, uno se ve forzado a asumir una doble imitación en la que *A* imita a *B* y *B* imita a *A*. Los objetos pueden surgir de este mecanismo conforme a la lógica de la profecía autocumplida. *A* cree que *B* desea a *O*, da sus primeros pasos hacia *O*, significando así a *B* que *O* es deseable. Cuando *B* a su vez manifiesta su deseo, *A* tiene la evidencia de que no se equivocó. El objeto *O* no es un origen, ya que es de hecho un efecto, un suplemento; en el origen solamente está la repetición.

Un razonamiento similar se puede aplicar al mecanismo de la víctima. La víctima sacrificada —hecha sagrada— paga por los demás, es un chivo expiatorio. Pero el uso de esta expresión revela cierto conocimiento del mecanismo de victimización que la colectividad de perseguidores no puede tener, porque tenerlo significaría el fin de dicho mecanismo. McKenna hace ver que "la víctima es siempre un sustituto, un significante (signifier)" y agrega "una marca, en términos derridianos, de una postergación". De hecho, el sacrificio colectivo calma la violencia a través de la violencia y es siempre una postergación de la violencia. Así, McKenna concluye: "Nada de lo que Derrida afirma a través de su crítica de

4 El término *différance* no existe en francés y fue introducido por Derrida para designar tanto una diferencia como una postergación.

los orígenes... constituye una prueba en contra de la hipótesis antropológica (Girard) de los orígenes humanos como enraizados en la dinámica de un deseo mimético".

Otra opción sería leer a Derrida a la luz de Girard. Esto es lo que Girard mismo hace en su texto citado aquí. Muestra cómo las "teorías del origen de la cultura no-científica", esto es, los mitos fundacionales, están incrustados en la lógica del suplemento por lo menos o más que la así llamada metafísica occidental. Girard explica esto usando su propia hipótesis antropológica: si los mitos acerca del origen se estructuran alrededor de una lógica paradójica, es porque cuentan un cuento que *efectivamente ocurrió*, pero que fue vivido en la ignorancia (*méconnaisance*). Una sociedad se hace unitaria por la vía de una exclusión: aquella de la víctima que se convierte en sagrada; sin dicha exclusión, la sociedad no existiría. El mito expresa al mismo tiempo la naturaleza interna y externa, el carácter dispensable e indispensable, y el infinito bien y el infinito mal de la víctima sacralizada. Lo último *es* el suplemento. La lógica del suplemento es la lógica de una narrativa que habla de un evento real con distorsiones que no son azarosas, sino bien definidas y reguladas. Este suplemento se encuentra en el corazón de todo pensamiento religioso y, debido a que lo religioso subsiste en lo filosófico, la filosofía también está debilitada por el suplemento. "La lógica del suplemento debe ser mítica primero y filosófica en segundo lugar".

Es esta epistemología realista lo que Dumouchel y Livingston abordan en sus textos. Dumouchel comenta que la teoría de Girard es una teoría morfogenética: puede explicar la emergencia de nuevas formas. Los mecanismos miméticos pueden producir nuevas complejizaciones, lo simple puede producir lo complejo. El origen, según Girard, no contiene lo que saldrá de él: no es una esencia. La crítica derridiana no tiene ningún efecto sobre la teoría girardiana dado que, como dice Derrida, uno no puede criticar la metafísica sino con conceptos metafísicos. Y como dice Dumouchel: "En Derrida está implícito un concepto de origen que no puede adaptarse a su misma crítica del origen". La teoría de Girard del origen de la cultura le permite a uno afirmar y concebir en forma

precisa un origen que no es una esencia.

En igual dirección, Livingston también afirma que es posible y deseable despejar al pensamiento girardiano de todo remanente metafísico. Para lograr esto es necesario colocar a Girard en un "marco naturalista" que es, según Livingston, lo que Girard mismo quiere hacer: "Girard busca evitar el tipo de circularidad que un desconstruccionista espera encontrar en cualquier teoría acerca del origen de la representación". Para hacer esto posible, uno debe encontrar la raíz de la mimesis humana en la mimesis animal, considerar que "esta mimesis tiene un origen: ha evolucionado o emergido como una realidad natural dentro del universo natural". Esto implica la existencia de una mimesis puramente instintiva "una forma inmediata y mecánica de mímica". Uno puede liberar a la mimesis original de todo aspecto representativo e intencional. En el origen no habría un "deseo" mimético, sino simplemente una mimesis de la apropiación. En este sentido, Girard puede explicar este mundo de representaciones que es la cultura, a partir de una hipótesis donde las representaciones no juegan ningún papel.

Si es posible leer la teoría de Girard por un lado en el sentido de la desconstrucción y por otro a la luz de una epistemología realista/positivista, es a causa de un aspecto específico de su configuración, a saber, el hecho de que el origen que postula es supuestamente un evento real, pero que solamente puede producirse porque su realidad es méconnue, interpretada erróneamente por quienes son sus actores. La teoría misma explica porque el origen es real pero inaccesible, siendo uno de esos extraños objetos que no existen salvo en la ignorancia. (Véase la Fig. 4, página siguiente)

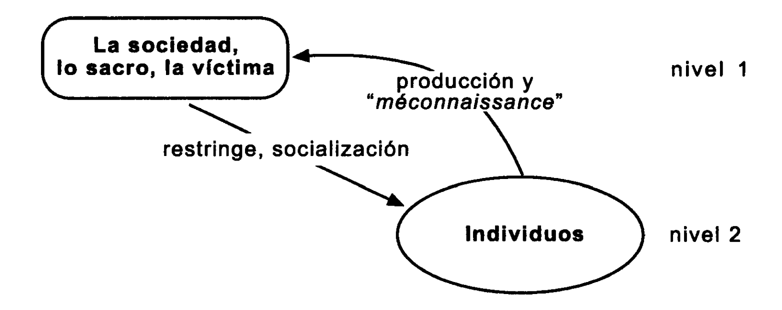


Figura 4

En consecuencia, nosotros sostenemos que la teoría de Girard y otras teorías acerca del origen, como veremos en breve, requieren un marco filosófico más adecuado. Esto sería una epistemología que no busque un fundamento último, real en el estilo de las ciencias duras, ni tampoco se satisfaga con el nihilismo de una permanente desconstrucción. Lo que se necesita es una posición intermedia, una metaposición que no necesite una fundamentación última y al mismo tiempo busque un mecanismo original para su fundamentación. La aparente contradicción de esta postura se resuelve, por cierto, exigiendo que el mecanismo de origen comparta las cualidades morfo-genéticas y paradójicas que hemos delineado hasta ahora, como se ejemplifica en los dos casos aparentemente no relacionados de organización celular y mimesis girardiana. Este es el núcleo de la interpretación que proponemos como una visión fresca frente a la cuestión del origen. Pero nos estamos adelantando un poco a nosotros mismos. Volvamos ahora a considerar el contenido de los otros casos.

La pregunta porque clarifica de manera ejemplar los temas fundamentales que hemos estado discutiendo. De hecho, la pregunta por lo que otorga valor al signo (token) monetario ha sido objeto de las más contradictorias interpretaciones. Al interior de este espectro encontramos nuevamente los dos extremos: el de la

excavación en un evento original real, absoluto y fundamental; y por otro lado, la desconstrucción hasta un completo nihilismo. La lectura de Orléan acerca de las teorías económicas modernas apunta precisamente en la dirección de una vía intermedia de autofundación.

Durante mucho tiempo, la economía del dinero ha estado dominada por la tradición metalista, la cual sostiene que "la aceptación fundacional del dinero como medio de intercambio debe entenderse como una creencia en el valor intrínseco del dinero mismo". Existe un bien determinado que tiene un valor "verdadero", por lo general referido al oro, y el dinero debe su valor al hecho de que contiene o al menos está respaldado por este valor real del oro.

Pocas personas mantienen hoy en día esta postura a la luz de la evidencia de que el papel moneda —llamado todavía dinero fi*at* — tiene valor incluso cuando carece de valor intrínseco y no es convertible. Es esta evidencia lo que explica por qué tanto el estructuralismo como el desconstructivismo han utilizado abundantes metáforas monetaristas para desconstruir la concepción metafísica de un signo que corresponde a un trascendental portador de significados. Estos desmitificadores modernos y posmodernos se apresuran a apuntar al papel moneda como un ejemplo de significación: signo meramente "autónomo", sin un referente, un signo de un signo, una copia de una copia, un simulacro. El tesoro central, el cual debería consagrar el valor del dinero está vacío y esto debe proclamarse fuerte y claro.

Orléan muestra cómo es posible encontrar otra alternativa en esta oposición estéril. Los orígenes del dinero no residen ni en una fundamentación última ni en un tiempo inaccesible, sino en un m*ecanismo* de autoorganización y auto-exteriorización. Este mecanismo de autoiniciación es el mismo que utiliza la sociedad para proyectarse a sí misma como si tuviera existencia fuera de sí misma, por así decirlo.

Una gran originalidad en el enfoque de Orléan es su focalización sobre la continuidad entre el dinero primitivo y el dinero

moderno, ya que "ambos son expresiones del todo social como una entidad separada". De este modo, re-formula el análisis de antropólogos como Barraud, de Coppet, Itéanu y Jamous cuando dicen: "no existe dinero en ausencia de un orden trascendente que le otorgue la cualidad de materialización de la totalidad social".

Esta realidad común hace posible distinguir al dinero moderno del dinero primitivo. Para las sociedades arcaicas, lo que realiza y "materializa" el movimiento a través del cual la sociedad se exterioriza es lo sacro (y Orléan establece aquí un vínculo con la antropología de Girard). Existe una gran riqueza de observaciones antropológicas sobre esto, y Orléans describe algunas de ellas, mostrando que el dinero primitivo está indisolublemente asociado a lo sacro. Las sociedades modernas no tienen acceso al mismo mecanismo de exteriorización, y no nos sorprende ver que las teorías monetarias ultra-individualistas, dado que niegan la trascendencia de lo social con respecto a los individuos ¡son llevadas a negar la realidad del dinero mismo!

Orléan favorece una visión monetarista particular, la de Keynes, porque piensa que deja en claro cómo un proceso de autoexternalización es posible cuando se lo considera en un marco individualista. El dinero y los valores individualistas pueden entonces conciliarse. La realidad del dinero toma la forma de convenciones, siempre contingentes en parte y sin basarse en alguna razón o naturaleza última, pero sin las cuales la sociedad no existiría de ninguna manera. Esta actividad social conducente a que dicha sociedad vea una parte de sus propios hechos como exteriores a sí misma —el dinero adquiriendo un valor que parece ser dado en otro lugar— introduce entonces una cierta opacidad que estaría ausente si hubiera algún tipo de fundamentación última. La naturaleza de este proceso de convenciones aparece entonces como otra instancia de una red altamente interactiva de agentes que encuentran mutua satisfacción por su actuar simultáneo. Esta solución endógena adquiere, a la vista de los actores, un valor externo, creando así precisamente la lógica del suplemento que ya hemos encontrado, como se muestra en la Figura 5.

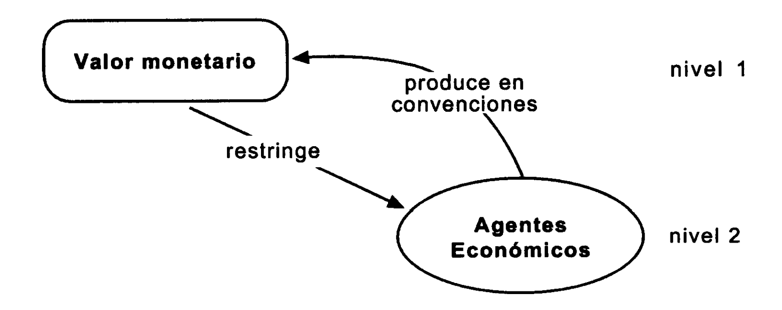


Figura 5

Jean-Joseph Goux destaca la persistencia de algunas características del dinero primitivo en el dinero moderno, comenzando por su dimensión de sacrificio: no existe ningún valor o significado estándar sin algún tipo de exclusión radical similar a un sacrificio, incluso en el caso del dinero moderno. Para ser exactos, como ya lo demuestra el caso griego, la modernidad introduce una ruptura fundamental que puede explicar al mismo tiempo el nacimiento del dinero moderno, el surgimiento del individualismo y la emergencia del ideal de la autonomía o auto-instauración de la sociedad. El Ser visto por los modernos (y ya por los griegos) implica una dimensión de trascendencia (y por ende de sacrificio) pero en la forma de auto-trascendencia (auto-sacrificio). El individualismo tiende a internalizar este movimiento más y más en el sujeto individual mismo. En el ámbito monetario, la mejor ilustración de esta des-trascendentalización paradójica es la circulación del estándar otrora trascendente del valor como un medio de intercambio, un medio de circulación en el mercado profano.

La evolución y la pregunta de cómo lo viviente ha llegado a ser como hoy lo encontramos es, obviamente, otro tema central del probema de los orígenes (ref. Stuart Kauffman, Brian Goodwin, Susan Oyama, John Dupré y Daniel Brooks). Debe decirse desde un principio que el análisis aquí está explícitamente hecho en el sector crítico de la biología evolutiva actual. De hecho, aunque la

noción de adaptación es la pieza central de gran parte de la biología evolutiva reciente, en años recientes han surgido muchas críticas de este supuesto programa adaptacionista, dando lugar a una revisión completa de lo que había sido hasta hace poco un punto de vista uniforme[5].

La ortodoxia que está siendo revisada hoy en día es la teoría de la evolución orgánica en su formulación neo-darwiniana, la cual es relativamente fácil de expresar en forma sucinta. Esta herencia puede resumirse en tres puntos básicos:

(1) La evolución ocurre como una modificación gradual de los organismos por descendencia; es decir, hay reproducción con herencia.

(2) Este material hereditario se diversifica constantemente (mutación, recombinación).

(3) Existe un mecanismo central para explicar *cómo* se producen estas modificaciones: el mecanismo de la selección natural. Este mecanismo opera a través de la selección de los diseños (fenotipos) que se relacionan con el medio ambiente actual de manera más eficiente.

Este darwinismo clásico se convirtió en neo-darwinismo durante la década de los treinta como resultado de la supuesta 'síntesis moderna' entre las ideas barwinistas basadas en la zoología,

5 Ver en particular: S.J. Gould. "Darwinism and the expansion of evolutionary theory" *Science* 216 (1982): 380-387; y S.J. Gould y R. Lewontin, "The Spandrels of San Marco and the Panglossian Paradigm: a critique of the adaptationist programme", *Proceedings of the Royal Society of London* 205 (1979): 581-598. Para una información más general ver Eliot Sober *The Nature of Selection* (Cambridge, Mass: MIT Press, 1984); M. Ho y P. Saunders, *Beyond Neo-Darwinism* (Nueva York: Academic Press, 1984); y J. Endler, 'The newer synthesis' Some conceptual problems in evolutionary biology,' *Oxford Surveys in Evolutionary Biology* 3 (1986) 224-243. Para una *defensa* reciente del neo-darwinismo a la luz de estos diversos cuestionamientos, ver: M. Hecht y A. Hoffman, 'Why not neo-Darwinism? A critique of paleobiological challenges?' *Oxford Surveys in Evolutionary Biology*, 3 (1986) 1-47, Nuestra discusión en esta sección también debe mucho a M. Piatelli-Palmarini, 'Evolution, selection and cognition' en E. Quagliariello, G. Bernardini y A. Ullman (editores), *From Enzyme Adaptation to NaturalPhilosophy* (Amsterdam: Elsevier, 1987), que explora temas similares, aunque en el contexto de una defensa del cognitivismo.

botánica y sistémica, por un lado, y el creciente conocimiento de la genética celular y de las poblaciones, por el otro. Esta síntesis estableció el punto de vista básico de que las modificaciones se producen a través de pequeños cambios en las características de los organismos que son especificadas por unidades heredables, los genes. La configuración genética responsable por los conjuntos de características conduce a tasas de reproducción diferenciales y, por lo tanto, a cambios en la configuración genética de una población animal a través de las generaciones. La evolución es simplemente la totalidad de estos cambios genéticos en poblaciones entrecruzadas. El ritmo y velocidad de evolución se miden por los cambios en la idoneidad de los genes; así, es posible dar una base cuantitativa para la adaptación visible de los animales a los ambientes en los que viven. Estos son conceptos, por cierto, con los cuales todos estamos familiarizados. Sin embargo, para hacer justicia a sus múltiples roles científicos, debemos clarificarlos un poco más.

Consideremos el concepto de adaptación. Su definición más intuitiva la señala como una forma de d*iseño* o construcción que coincide óptimamente (o por lo menos muy bien) con alguna situación física. Por ejemplo, las aletas de los peces son muy adecuadas para un ambiente acuático, en tanto que el casco ungular se adapta bien para correr en las praderas. Aun cuando este concepto de adaptación es bastante popular, la mayoría de los teóricos evolucionistas profesionales no la interpretan de este modo. Para ellos, la adaptación ha llegado a referirse específicamente al *proceso* vinculado a la reproducción y la supervivencia, es decir, con la adaptación. Este proceso es —eso suponemos— lo que explica el aparente grado de diseño adaptativo que se observa en la naturaleza.

Para hacer que esta idea de adaptación haga su trabajo teórico, sin embargo, es necesario analizar de alguna manera el grado de adaptación de los organismos. Es aquí donde aparece la noción de *adecuación óptima*. Desde el punto de vista del grado de adaptación, la tarea de la evolución consiste en encontrar estrategias heredables, es decir, conjuntos de genes interrelacionados que serán más o menos capaces de contribuir a la reproducción diferencial. Cuando un gen

cambia en aras de un mejor cumplimiento de esta tarea, mejora su adecuación. Esta idea de adecuación muchas veces se formula como una medida de abundancia. Por lo general se la interpreta como una medida de abundancia individual, es decir, como una medida del excedente de descendencia logrado, pero puede también interpretarse como una medida de abundancia de la población, es decir, como el efecto de los genes en la tasa de crecimiento de una población.

Se ha hecho cada vez más claro, sin embargo, que esta forma de medir la adecuación como abundancia tiene una cantidad de dificultades conceptuales y empíricas. En primer lugar, en la mayoría de los grupos animales, el éxito reproductivo depende de los encuentros sexuales con otros individuos. En segundo lugar, dado que los efectos de cualquier gen están siempre entrelazados con una multitud de otros genes, no siempre es posible diferenciar los efectos de los genes individuales. En tercer lugar, el medio ambiente, en el cual se supone que los genes deben expresarse, es enormemente variado y dependiente del tiempo. Por último, este medio ambiente debe verse en el contexto del ciclo de vida completo y la ecología de un animal. La adecuación puede además tomarse como una medida de *persistencia*. En este caso, la adecuación mide la probabilidad de la permanencia reproductiva en el tiempo. Lo que se optimiza no es la cantidad de la descendencia sino la probabilidad de la extinción. Claramente, este enfoque es más sensible a efectos en el largo plazo, y por lo tanto constituye una mejora sobre el punto de vista de la idoneidad como abundancia. Asimismo, sin embargo, plantea enormes problemas en términos de medición. Prácticamente todos estos puntos se abordan en los siguientes textos.

Armada de estos refinamientos, en las últimas décadas la ortodoxia dominante en el pensamiento evolucionista vio la evolución como un 'campo de fuerzas'[6]. Las presiones selectivas (la metáfora física es apropiada) actúan sobre la diversidad genética

6 Este término se tomó de Eliot Sober, *The Nature of Selection* (op. cit.)

de una población, produciendo cambios en el tiempo de acuerdo con una optimización del potencial de idoneidad. La postura adaptacionista o darwiniana proviene de entender este proceso de selección natural como el factor central en la evolución orgánica. En otras palabras, la teoría evolucionista ortodoxa no niega que haya varios otros factores que operan en la evolución; simplemente baja el perfil de su importancia y busca explicar los fenómenos observables únicamente sobre la base de la optimización de la idoneidad. El origen de lo vivo y su diversidad se refiere fundacionalmente a este referente externo que es el paisaje de una idoneidad en constante mejora.

El desafío a esta ortodoxia adopta una forma que a esta altura debería sernos familiar. Dicho en términos simples: las características intrínsecas de la vida misma son la fuerza central de la evolución, dando forma a lo que es o no es adecuado. Así, lo Otro es visto bajo esta perspectiva no como origen, sino como resultado. Una vez más tenemos la lógica de la *différance* en acción. Para esta re-lectura de la evolución, los biólogos modernos colocan en un primer plano varios factores, todos relativos a la riqueza de los organismos como redes (de características, de genes, de comportamientos). Son estas cualidades auto-organizativas las que constituyen la flecha en dirección contraria desde lo que parece ser jerárquicamente dependiente a lo constitutivo y causal. Una lista parcial de estos factores incluye a la interdependencia genética (pleiotropía), restricciones al desarrollo, derivación genética, estasis evolutiva, y los diversos niveles de las unidades de selección. Como dijo Richard Lewontin en una crítica reciente a la posición clásica: "no es que estos fenómenos [es decir, las restricciones al desarrollo, pleiotropía, etc.] no sean mencionados, sino que son claramente desviaciones del evento principal, la ascensión al Monte Idoneidad por Sir Ronald Fisher y sus fieles sherpas"[7]. Los biólogos evolucionistas, de manera creciente, se han involucrado en un movimiento que se aleja del Monte Idoneidad hacia una nueva teoría,

7 R. Lewontin, 'A natural selection: Review of J.M. Smith's *Evolutionary Genetics*; *Nature* 339 (1989):107.

más amplia y hasta ahora incompletamente formulada[8]. A Alfred Russel Wallace le gustaba decir que "Nada en la Naturaleza **no es** útil"[9]. El sentir de nuestra cultura aún concuerda con los argumentos que apelan a la parsimonia de la naturaleza y sus diseños óptimos, y muchos científicos siguen utilizándolos sin pensarlo dos veces. La lista de temas entrelazados que acabamos de discutir, sin embargo, señala otra direcciones.

Todos estos temas se abordan en el texto de Kauffman y en los comentarios que se refieren a él. El enfoque de Kauffman es muy ingenioso, ya que se fija en un claro caso de adaptación, un espacio de pequeñas proteínas, y procede a complejizarlo hasta el punto en que los efectos opuestos y paradójicos de los factores internos comienzan a aparecer. Aunque sus argumentos de ninguna manera prueban una gran teoría unificada, son algunos de los mejores que se han producido hasta el momento en términos de su claridad y rigor para explorar lo que una teoría pos-darwinista podría parecer.

Brian Goodwin se suma a los aspectos tratados por Kauffman enfocando otro que toma prestado de una disciplina central —aunque muy descuidada— en la evolución: el desarrollo. Su argumento principal es que las restricciones al desarrollo son tan importantes y penetrantes que una gramática del organismo es lo mismo que uno ve en la historia de la biosfera. Daniel Brooks contribuye con un ángulo original a la cuestión de la revisión del neo-darwinismo, mediante el examen de la circulación de energía en la transformación de la materia y de este modo, entender los conceptos de entropía e "información" como un marco común para los procesos de desarrollo y evolutivos. Susan Oyama y John Dupré, por otro lado, continúan el debate en un tono más filosófico. El punto de Dupré es evaluar

8 Un interesante ejemplo de este ánimo revisionista es el estudio crítico del ejemplo clásico de melanismo industrial en las polillas como un caso de estudio de la selección natural. De acuerdo con D. Lambert, C. Millar y T. Hughes, 'On the classic case of natural selection', *Biology Forum* 79 (1986): 11-49, este ejemplo puede transformarse en un estudio clásico contra el neo-darwinismo tomando en cuenta una cantidad substancial de literatura existente que ha sido ignorada.

9 H. Clemens, *Alfred R. Wallace: Biologists and Social Reformer* (Londres: Hutchinson, 1983).

los argumentos de Kauffman y examinar si acaso no están dependiendo todavía demasiado de algún marco adaptacionista. Argumenta en favor de una teoría evolutiva multifacética con espacio para varios mecanismos de niveles múltiples. Oyama aborda la importante cuestión de la oposición naturaleza-nutrición a la luz de las teorías neo-darwinistas. Según dice, esta oposición es siempre indicativa de una severa línea demarcatoria entre un exterior y un interior. Cuando estos dos se interrelacionan —como lo exige la lógica del suplemento— la oposición desaparece en otro paisaje conceptual.

Para concluir, volvamos sobre algunas ideas centrales: el centro de la cuestión es que la explicación de una regularidad biológica observada, como un calce o como una correspondencia óptima con las dimensiones dadas del medio ambiente, parece cada vez menos sostenible sobre bases lógicas y empíricas. Parte de la dificultad para ir más allá de la estructura neo-darwinista es determinar qué hacer una vez que se abandona la idea de la selección natural como explicación principal, lo que equivale a decir que las estructuras, mecanismos, características o disposiciones no pueden explicarse sobre la base de su contribución al valor de supervivencia. Se siente uno tentado a preguntar: entonces ¿están allí las cosas sin razón alguna? La tarea de la biología evolutiva pos-darwiniana es cambiar la geografía lógica del debate estudiando las complejas *relaciones circulares* de congruencia entre los ítemes por explicar. Esto nos coloca de lleno en el mismo tipo de lógica evocado antes, como se bosqueja en la Figura 6.

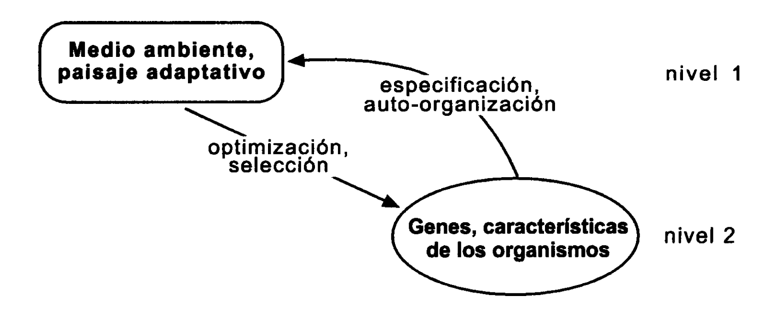


Figura 6

De acuerdo con la sabiduría tradicional, el medio ambiente en el cual los organismos evolucionan es algo dado, fijo y único. Aquí nuevamente encontramos la idea de que los organismos llegan a un medio como caídos del cielo. Este punto de vista se refina un poco cuando se concede la posibilidad de pequeños cambios en el medio, concesión que ya era empíricamente familiar a Darwin. Un ambiente en movimiento de este tipo proporciona las presiones selectivas que forman la columna vertebral de la teoría evolutiva neo-darwiniana. Al avanzar hacia una visión revisada de la evolución, sin embargo, introducimos un factor adicional: redefinimos las presiones selectivas y les asignamos el nuevo papel de restricciones amplias que deben ser satisfechas. El punto crucial aquí es que no sostenemos la noción de un ambiente autónomo, dado de antemano, sino que permitimos que se disuelva en el entorno en favor de los así llamados "factores intrínsecos", Lo que enfatizamos es que la noción misma de lo q*ue es un medio ambiente no puede separarse de lo que los organismos son y hacen.* Este punto ha sido defendido de manera bastante elocuente por Richard Lewontin:

> *"...el organismo y el medio ambiente no se determinan por separados. El medio no es una estructura impuesta a los seres vivos desde afuera; en realidad, es una creación de esos seres. El medio no es un proceso autónomo, sino un reflejo de la biología de las especies. Así como no existe un organismo sin un medio, tampoco hay un medio sin un organismo"*[10].

El punto clave es, pues, que las especies generan y especifican su propio campo de problemas por 'resolver' por 'satisfacción'; este campo no existe 'afuera', en un medio que operaría como cancha de aterrizaje para organismos que caen o se lanzan en paracaídas sobre el mundo. En lugar de eso, sucede que los seres vivientes y sus entornos se posicionan con relación a sí mismos a través de una esp*ecificación mutua* o *co-determinación.* Así, lo que describimos como regularidades ambientales no son características externas que han sido internalizadas, como asumen tanto el representacionismo

10 Richard Lewontin. 'The organism as the subject and object of evolution', *Scientia* 118 (1983): 63-82.

como el adaptacionismo. Las regularidades ambientales son el resultado de una historia conjunta, una congruencia que se despliega a partir de una larga historia de co-determinación. En palabras de Lewontin, el organismo es, a la vez, el sujeto y el objeto de la evolución[11].

El tópico final que prometimos abordar fue el origen del contenido cognitivo del significado del mundo como lo percibimos. Más específicamente, la discusión está centrada en las capacidades cognitivas básicas: percepción y acción, más que con capacidades así llamadas superiores tales como el lenguaje y la memoria.

En mi trabajo he propuesto visualizar la ciencia cognitiva en tres etapas sucesivas. Estas tres etapas corresponden a tres momentos sucesivos: cada uno indica un importante cambio en el marco teórico de la ciencia cognitiva. El centro o núcleo de las ciencias cognitivas se conoce generalmente como cogn*itivismo*[12]. La herramienta central y metáfora guía del cognitivismo es la computadora digital, entendida como un dispositivo físico construido de tal manera que un conjunto particular de sus cambios físicos pueda interpretarse como computaciones. Una computación es una operación realizada o llevada a cabo en símbolos, es decir, en elementos que *representan* a aquello cuyo lugar ocupan. Simplificando un poco, podemos decir que el cognitivismo consiste en la hipótesis de que el entendimiento —incluyendo el humano— es la manipulación de símbolos a la manera de los computadores digitales. En otras palabras, el entendimiento es una *representación mental*: la mente estaría operando a través de la manipulación de símbolos que representan características del mundo, o representan al mundo como si fuera de cierta manera. De acuerdo con esta

---

11 Richard Lewontin. 'The organism as the subject and object of evolution', *Scientia* **118** (1983): 63-82.

12 Esta designación es justificada por John Haugeland, 'The Nature and Plausibility of Cognitivism', reimpreso en *Mind design: Philosophy, Psychology, Artificial Intelligence*, John Haugeland (ed.) (Montgomer, Vt. Bradford Books, 1981). Ocasionalmente, el cognitivismo es descrito como el 'paradigma simbólico' o el 'enfoque computacional'. Tomamos estas designaciones como sinónimos para efectos de lo que nos interesa aquí.

hipótesis cognitivista, el estudio del entendimiento como una representación de la mente proporciona el campo propio de la ciencia cognitiva, campo que se reconoce como independiente de la neurobiología por un extremo y de la sociología y de antropología por el otro.

El cognitivismo tiene la virtud de ser un programa de investigación bien definido, que incluye prestigiosas instituciones, revistas, tecnologías aplicadas e intereses comerciales internacionales. Nos referimos a él como el centro o núcleo de la ciencia cognitiva porque domina la investigación a tal grado que muchas veces simplemente se considera que *es* la ciencia cognitiva misma. En años recientes, sin embargo, han aparecido varios enfoques alternativos sobre la cognición. Estos enfoques difieren del cognitivismo en dos líneas básicas:

(1) Una crítica al procesamiento de símbolos como vehículo apropiado para las representaciones.

(2) Una crítica de lo adecuado de la idea de representación como punto de Arquímedes de la ciencia cognitiva.

La primera alternativa, que podríamos llamar "emergencia" es típicamente llamada conexionismo. Este nombre deriva de la idea de que muchas tareas cognitivas (por ejemplo, la visión y la memoria) parecen ser manejadas de mejor manera por sistemas consistentes en muchos componentes simples, los cuales, conectados a través de reglas apropiadas, dan lugar a un comportamiento global correspondiente a la tarea deseada. El procesamiento de símbolos cognitivista, por el contrario, es localizado: las operaciones sobre los símbolos pueden especificarse usando solamente la forma física de los símbolos, no su significado. Por cierto, es esta característica de los símbolos lo que le permite a uno construir un dispositivo físico para manipularlos. La desventaja es que la pérdida de cualquier parte de los símbolos o las reglas para su manipulación da como resultado serias disfunciones. Los modelos conexionistas, por el contrario, generalmente cambian el procesamiento local, simbolizado por operaciones distribuidas, es decir, operaciones que se extienden

sobre una red completa de componentes y, por lo tanto, dan como resultado la emergencia de propiedades globales que son adaptables a las disfunciones locales. Para los conexionistas, una representación consiste en la correspondencia entre dicho estado global emergente y propiedades del mundo; no se trata de una función de símbolos particulares.

La segunda alternativa nace de una insatisfacción más profunda que la relacionada con la búsqueda conexionista de alternativas para el procesamiento simbólico. Cuestiona la centralidad de la noción de que el entendimiento es fundamentalmente representación. Detrás de esta noción hay dos suposiciones fundamentales. La primera es que habitamos un mundo con propiedades particulares, tales como largo, color, movimiento, sonido, etc. La segunda suposición es que 'recogemos' o 'recuperamos' estas propiedades representándolas internamente. Estas dos suposiciones implican un compromiso fuerte, muchas veces tácito y más allá de todo cuestionamiento con el realismo u objetivismo, en torno a cómo es el mundo y cómo llegamos a conocerlo.

Existen, sin embargo, muchas maneras en que el mundo es —de hecho incluso muchos distintos mundos experienciales—, dependiendo de la estructura del ser involucrado y los tipos de distinciones que es capaz de hacer. Más aún, si restringiéramos nuestra atención al entendimiento humano, existen muchas maneras en las que puede explicarse el mundo[13]. Esta convicción no-objetivista crece lentamente en el estudio de la cognición. Por lo pronto, sin embargo, esta orientación alternativa no tiene un nombre establecido, ya que es más bien un paraguas que cubre a un grupo relativamente pequeño de personas que trabajan en diversos campos. Proponemos como nombre el término en*activo* para enfatizar la creciente convicción de que el entendimiento no es la representación de un mundo pre-dado sino la puesta en escena o la gestación de un mundo sobre la base de la historia y la diversidad de acciones efectivas que un ser puede llevar a cabo. El enfoque enactivo

13 Ver Nelson Goodman, Ways of Worldmaking (Cambridge/Indianapolis: Hacket Publishing Company, 1978).

considera seriamente, entonces, la crítica filosófica a la idea de que la mente es un espejo de la naturaleza, pero va más allá al abordar el tema desde el interior del ámbito de la ciencia[14].

A estas alturas el lector se habrá dado cuenta de que cada alternativa para el estudio de la cognición ensancha el campo de la ciencia cognitiva. Entonces, en cada etapa enfocamos algo que antes habíamos considerado por su valor nominal, como un símbolo o representación. Por último, en la periferia nos preocupamos explícitamente del origen de las regularidades que normalmente parecen estables y fijas y, por lo tanto, receptivas a un tratamiento simbólico o representativo. Habiendo llegado a este punto, podemos avanzar en la dirección opuesta, desde la periferia hacia el centro poniendo entre paréntesis la preocupación por el origen de estas regularidades, en orden a considerarlas simplemente en su valor nominal.

El trabajo de Christine Skarda complementa nuestra argumentación desde otro ángulo. La base de sus comentarios es el importante trabajo de Walter Freeman, quien ha introducido la noción de que la constitución dinámica y las propiedades de las redes neuronales son en sí mismas lo suficientemente ricas como para poner en tela de juicio la idea de que el cerebro funciona a través de representaciones en el sentido clásico. La dinámica de los conjuntos neuronales es en la actualidad un foco de investigación muy activo, y Skarda sostiene que la distinción dinámica versus conexionista/simbólica es la más relevante para el origen de la percepción.

Resulta interesante mencionar cómo la percepción de los procesos evolutivos, como se describe arriba, puede de hecho ser entendida por el concepto representacionista de que existe una correspondencia entre organismo y medio que es dada por las restricciones optimizadoras de supervivencia y reproducción. Dicho llanamente, el representacionismo en la ciencia cognitiva es exactamente ho*mólogo* al adaptacionismo en la teoría evolucionista, ya que el carácter óptimo juega el mismo papel central en cada dominio. Se

---

14 Ver Richard Rorty, *Philosophy and the Mirror of Nature* (Princeton University Press, 1979).

desprende, entonces, que cualquier evidencia que debilite el punto de vista adaptacionista genera *ipso facto* dificultades para el enfoque representacionista del entendimiento.

Ahora bien, los científicos cognitivos han sido guiados inexorablemente por las exigencias de sus investigaciones al estudio de las sub-redes que actúan a nivel local. Estas redes interactúan unas con otras en redes entretejidas, formando 'sociedades' o 'agentes' para usar terminología de Minsky. Debería ser claro, a partir de nuestra lista de actuales problemas, que los teóricos evolucionistas han llegado de manera independiente a prácticamente las mismas conclusiones. La supervivencia y la reproducción como restricciones son demasiado débiles como para proporcionar una descripción acerca de cómo se desarrollan y cambian las estructuras. Asimismo, ningún esquema de optimidad g*lobal* parece ser suficiente para explicar los procesos evolutivos. Existen, por cierto, 'agentes' genéticos locales para, digamos, el consumo de oxígenos y el crecimiento de las plumas, los cuales pueden medirse de acuerdo con alguna escala comparativa en donde se pueda buscar la optimalidad, pero ninguna escala en particular servirá para todos los procesos. Por ende, nos encontramos de vuelta en nuestra conocida lógica del suplemento para los fenómenos cognwitivos, según se ilustra en la Figura 7.

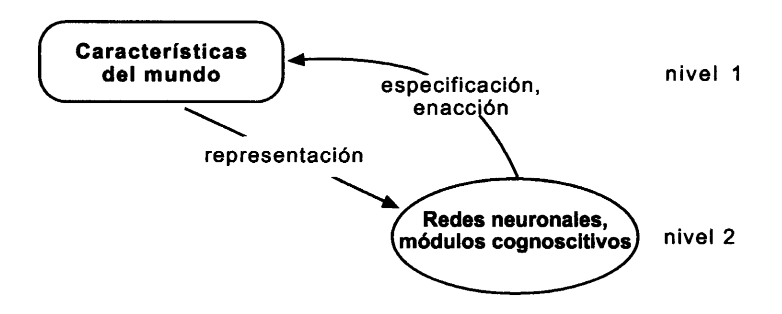


Figura 7

El tema central de la instancia biológica, la evolución y el significado cognitivo puede ponerse en forma de una analogía[15]. Juan necesita un traje· En un mundo plenamente simbólico y representacionista, Juan va donde su sastre que le toma las medidas y le hace un traje exactamente a la medida. Existe, sin embargo, otra posibilidad obvia: una que no exige tanto del medio . Juan busca en las grandes tiendas y elige un traje que le queda bien entre los varios que hay disponibles. Aunque estos no le quedan *perfectos*, son lo suficientemente buenos y elige el traje óptimo en términos de talla y gusto. Tenemos aquí una buena alternativa seleccionista que aplica algunos criterios óptimos de idoneidad. La analogía admite, sin embargo, un mayor refinamiento. Juan, como cualquier ser humano, no puede comprar un traje como un hecho aislado del resto de lo que sucede en su vida. Al comprar un traje deberá considerar cómo su aspecto afectará la respuesta de su jefe en el trabajo, la posible respuesta de su novia, y también puede estar preocupado por factores políticos y económicos. De hecho, la decisión misma de comprar un traje no está dada desde un principio como un problema, sino que está constituida por la situación global de su vida. La elección final *satisface* algunas restricciones bastante poco definidas (por ejemplo andar bien vestido), pero no resuelve (menos aún de manera óptima) ninguna de las restricciones en particular.

Así, hemos revisado los casos de estudio más importantes. Todos están permeados por la tendencia a pensar en términos de lo que podríamos llamar un punto de referencia *exógeno* o punto fijo, o una colección de ellos. Estos deben ser impenetrables e inamovibles y el orden surge de ellos como una armonía preestablecida con ellos mismos, ya sea Dios o sus sustitutos terrenales: el oro, la optimización ambiental, la representación de características. Sin embargo, como hemos visto en repetidas ocasiones, tal fundamentación de un origen sobre

---

15 Para la mejor y más completa discusión técnica de este punto, ver G. Oster y S. Rocklin, Optimization models in evolutionary biology', *Lectures in Mathematical Life Sciences*, vol. 11 (Rhode Island: American Mathematical Society, 1979), páginas 21-88. Para una discusión reciente del tema, ver J. Dupré (editor) *The latest on the best* (Cambridge, Mass.: MIT Press, 1987).

la base de un punto fijo exógeno es propensa, o más bien, está infectada de un corrosivo derridiano. Nunca se da el caso de que tales externalidades puedan afirmarse en otra cosa que el suplemento del orden que supuestamente ellas mismas engendran. Se da entonces la reacción violenta hacia el otro extremo: Dios ha muerto, el mundo no es sino un caos eterno, sin belleza ni origen, sin significado. No hay nada más que interpretaciones de interpretaciones *ad infinitum*. "No hay hechos, sólo interpretaciones; nuestro discurso es nuestro 'nuevo infinito'", las citas son todas de Nietzsche[16]. (De hecho, el pensamiento nietzschiano está más enraizado en la actualidad de lo que uno pudiera pensar).

Pero esa es la interesante paradoja. Entre la tendencia a una concepción exógena de orden y el nihilismo existe una complicidad, a pesar de sus posiciones aparentemente contradictorias. Para ir de la primera a la segunda, solamente necesitamos una dosis de lógica del suplemento, para librarnos de lo que parece ser un punto fijo exterior al sistema. El paso desde lo segundo a lo primero es menos obvio y es donde falla el nihilismo: ver que existen principios de surgimientos endógenos, interdependencias trabajadas que son capaces de dar lugar a puntos fijos *endógenos*. Esto es lo que hemos rastreado aquí para los casos del orden social, dinero, evolución y entendimiento perceptual. Una vez que sacamos sus respectivos y aparentemente exógenos puntos fijos, podemos ver el establecimiento de un círculo completo de autorreferencia que da lugar a una constitución endógena no-arbitraria. Para este punto fijo endógeno, existe orden: sociedad, valor, especies y objetos en el mundo. Sin embargo, a diferencia de la argumentación del punto de vista exógeno de los orígenes, dichas externalizaciones no se basan en otra cosa que en los agentes constitutivos mismos y los

16 Esta analogía fue propuesta por primera vez en G. Edelman y W. Gall, 'The Antibody Problem', *Annual Review of Biochemistry* 38 (1979): 699-766. También fue usada por M. Piatelli-Palmarini, 'Evolution, selection and cognition' en E. Quagliarello, G. Bernardi y A. Ullman (editores), *From Enzyme Adaptation to Natural Philosophy*. Usamos la analogía aquí con una extensión que no coincide con la intención de ninguno de estos autores.

procesos particulares que los entrelazan. Todos se nutren de una lógica generativa fundamental[17]. Cuando este círculo creativo se rompe y aplana en una oposición unidimensional, recuperamos las tendencias antitéticas del orden exógenos versus desconstrucción (ver Figura 8.)

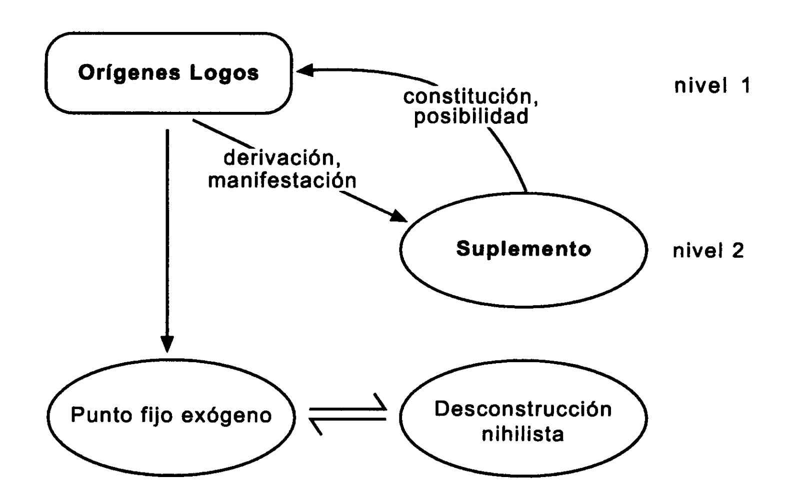


Figura 8

Así, proponemos que el orden y su origen pueden ser concebidos —no importa cuán tenue o incompletamente se haga— de una manera que escapa tanto a la tentación de un fundamento último como referencia exógena a una Razón o Verdad última, *como* a la tentación de la total ausencia de cualquier regularidad y la desconexión de toda forma de orden. La clave está en el descubrimiento, para cada caso y en cada campo —pero compartiendo una lógica común—, de un proceso morfogenético capaz de autofundamentarse y autodistinguirse. Un origen que no es ni inexistente o esquivo, ni tampoco una fundamentación última o una referencia absoluta.

---

17 C.f § 109 y 374 *Gay Science* y II § 133 *The Will to Power.*

IV

DISCUSIONES

RETRO-PERSPECTIVA

Orígenes de una idea

La noción de autopoiesis fue presentada por primera vez en español en un libro que apareció en la Editorial Universitaria, de Santiago de Chile, en los atribulados días de septiembre de 1973. El concepto (y, por ende, el libro) ha tenido una proyección y un impacto que ha perdurado hasta hoy, para mi gran regocijo. A los veinte años de su aparición, la Editorial nos solicitó, con Humberto Maturana, que escribiéramos extensos prefacios para una nueva edición. Opté por hacer una retrospectiva de los orígenes de la idea y de su contexto, que tiene su lugar natural en esta selección de ensayos.

Agradezco las circunstancias que permiten la reedición de este libro y la oportunidad de escribir este prefacio a 20 años de su primera edición. Es raro en la vida de un científico tener la ocasión, no sólo de producir ideas que se presienten como de envergadura, sino además de ver sus dificultades balbucientes y constatar cómo echan raíces en el pensamiento científico. Esto ha ocurrido con la noción de *autopoiesis y* sus vicisitudes, que este libro introducía por primera vez en 1973, y de las que he sido testigo de primera mano en el devenir científico internacional.

¿Qué hace que una idea como la *autopoiesis,* estrictamente una teoría de la organización celular, adquiera visibilidad y prominencia más allá de la biología profesional y sea capaz de afectar campos de saber lejanos? Mi respuesta es que en último

---

(Varela,F. (1996), Prefacio a la 2ª edición, *De Máquinas y Seres Vivos*, Editorial Universitaria, Santiago).

término sólo podemos comprender ese fenómeno porque la idea contiene un *trasfondo de sensibilidades históricas* de importancia con las cuales se alinea y resuena. Ese trasfondo de tendencias no aparece delineado sino en retrospectiva, porque las ideas, como la historia, son una posibilidad que se cultiva, no un determinismo mecánico. A esta distancia, la *autopoiesis* ocupa en mi opinión un lugar privilegiado por haber anunciado de manera clara y explícita una tendencia que hoy es ya una configuración de fuerzas en muchos dominios del quehacer cultural.

La tendencia a la que hago referencia, dicho rápidamente, es la desaparición del espacio intelectual y social que hace del conocer una representación mentalista y del hombre un agente racional. Es la desaparición de lo que Heidegger llama la época de la imagen del mundo y que puede también designarse como cartesianismo. Si la *autopoiesis* ha tenido influencia es porque supo alinearse con otro proyecto cuyo centro de interés es la *capacidad interpretativa del ser vivo* que concibe al hombre no como un agente que "descubre" el mundo, sino que lo constituye. Es lo que podemos llamar el giro ontológico de la modernidad, que hacia fines del siglo XX se perfila como un nuevo espacio de vida social y de pensamiento que ciertamente está cambiando progresivamente el rostro de la ciencia.

A lo largo es estas páginas intentaré un mayor desarrollo de lo que acabo de expresar. Son ideas que debemos tener como faros frente a nosotros, para hacer posible una lectura de los orígenes, la gestación y maduración de la idea de *autopoiesis.* Dicho de otra manera, la *autopoiesis* ocupa un lugar en una trama bastante más amplia que la de la biología, en la que aparece hoy en una posición privilegiada. Esa sintonía con una tendencia histórica, intuida más que sabida, constituye el fundamento central de este libro y es su trayectoria la que me propongo trazar.

Dejar una firma en un texto, más que una posesión personal, es un hito en un camino. Las ideas aparecen como movimientos de redes históricas en que los individuos son formados, más que ellos a las ideas. Así, Darwin tenía ya a Wallace que lo esperaba, y

a la Inglaterra victoriana como sustrato; Einstein solitario en su oficina de patentes suiza, dialogaba con Lorentz, y lo sostenía el mundo de la física germánica de fin de siglo; Crick conocía ya las ideas de Rose y Pauling al encontrar a Watson, y su estado de ánimo era propio del Cambridge de los años 50. Haciendo las diferencias y guardando las distancias que cabe, la historia de la autopoiesis también tiene sus antecedentes de resultados de donde surge y un sustrato peculiar que la nutre, en particular en las ideas de Maturana en los años 60. Pero más allá, fue Chile entero que jugó un rol fundamental en esta historia. Los nuevos científicos de Chile y América Latina tienen aquí material para la reflexión.

Escribir este prefacio es, insisto, un pliegue de la historia donde los hombres y las ideas viven porque somos más puntos de acumulación de las redes sociales en las que habitamos que voluntades o genios individuales. No se puede pretender aglutinar la densidad de acciones y conversaciones que nos constituyen en un relato personal necesariamente unidimensional. No pretendo que lo que digo aquí es una narrativa pretendidamente objetiva. Lo que ofrezco es, por primera vez, mi lectura, tentativa y abierta, de cómo surgió la noción de *autopoiesis*, y cuál ha sido su importancia y devenir. Cada una de las cosas que digo la he madurado largamente, y la creo honesta hasta donde puedo juzgar en mi conciencia como responsable de ser uno de los actores directos de esta creación, pero consciente de que no puedo considerarme poseedor de la verdad.

Para poder iluminar los temas de fondo necesito comenzar por lo que fueron las raíces de esta historia desde mi perspectiva personal. Paradójicamente, sólo a través de rescatar cómo los temas de fondo aparecieron en la especificidad de mi perspectiva es que puedo comunicar al lector la manera como esta invención encuentra un lugar en un horizonte más amplio.

Los años de incubación

Pertenezco a una generación de científicos chilenos que tuvimos el privilegio de ser jóvenes en uno de los momentos más creativos

de la comunidad científica chilena en la década de los años sesenta. Como adolescente tuve una vocación temprana por el trabajo intelectual y las ciencias biológicas me parecían sin dudas como mi norte. Al egresar de secundaria en 1963 opté por la Universidad Católica que anunciaba un innovativo programa de "Licenciatura en Ciencias Biológicas" consecutivo al tercer año de Medicina. Como alumno de medicina, conocí así a los primeros investigadores que me parecieron fascinantes, personajes como Luis Izquierdo, Juan Vial, Héctor Croxato y sobre todo Joaquín Luco, quien me contagió definitivamente con la pasión de la neurobiología. A poco andar en mi primer año, pedí a Vial que me admitiera como aprendiz en su laboratorio de Biología Celular. Me dio la llave de una pequeña puerta de su laboratorio que daba a calle Marcoleta, donde iba en horas libres a hacer cortes de nervios con tinción de mielina.

Juan Vial me dio también buenos consejos, incluyendo el de cambiarme a continuar mi formación en 1965 a la recién abierta Facultad de Ciencias de la Universidad de Chile. Fue un paso crucial, porque salí del universo de las carreras tradicionales para entrar de lleno en el universo de la formación científica exclusiva, hasta entonces desconocida en Chile. En unas salas prestadas en el último piso de la Escuela de Ingeniería de Beaucheff encontré mi caldo de cultivo: un pequeño grupo de jóvenes entusiasmados por la investigación en ciencias puras, y profesores-investigadores que enseñaban a los futuros científicos con pasión.

Aprendiz de neurobiólogo

El último de los consejos de Vial fue que intentara trabajar con Humberto Maturana, quien acababa de cambiarse de la Escuela de Medicina de la U. de Chile a la nueva Facultad de Ciencias. Un buen día de abril de 1966 fui a verlo a su laboratorio en el subterráneo de uno de los "dientes" de la nueva escuelaâMn calle Independencia. Me preguntó qué era lo que me interesaba y en mi entusiasmo de los veinte le dije sin vacilación: "¡El psiquismo en el universo!". Humberto sonrió y dijo: "Muchacho, has llegado al lugar adecuado...". Fue un día memorable y el comienzo de una relación que tendría

consecuencias para ambos. Maturana era ya un investigador de estatura en aquella época, conocido por sus trabajos de la fisiología de la visión en varios *papers* clásicos que había realizado en Harvard y en el MIT antes de volver a Chile[1]. En Chile continuaba trabajando en fisiología y anatomía de la retina de vertebrados.

Para continuar con mi aprendizaje en el oficio, Humberto me pidió que repitiera experimentos de registro eléctrico en el tectum óptico de la rana, lo que me llevó a profundizar los problemas de visión más a fondo de lo que nunca lo había hecho con ningún tema científico. Cuando dejé el laboratorio de Independencia para partir a Estados Unidos dos años después, había desarrollado la capacidad de generar mis primeras ideas de investigación. Se trataba de una hipótesis sobre el rol del tiempo en la forma de operar de la retina, que llevaba a unas predicciones experimentales que dieron origen a mi primer artículo científico[2].

La influencia de Maturana fue uno de los pilares que me dio los años de aprendizaje en Chile, y sobre la cual tendré más que decir. Pero es importante que esboce cómo hubo al menos otras dos corrientes de influencia que tuvieron y han seguido teniendo un enorme peso en mi historia intelectual. La primera, fue la filosofía y haber encontrado ciertas lecturas clave en esos años de formación. La segunda, fue el descubrir el mundo de la cibernética y la biología teórica.

La reflexión filosófica

En los años de liceo mis lecturas filosóficas eran tan apasionadas como azarosas, mezclando Aristóteles (en esas bellas ediciones de la *Revista de Occidente),* Ortega y Gasset, Sartre y Papini. En la Escuela de Medicina, un encuentro fructífero con Arturo Gaete

1 En particular véase el "clásico": Maturana, H.; Lettvin, J.; McCulloch, W.; Pitts, W., "Anatomy and physiology of vision in the frog", en J. *General Physiology,* 1960, 43: 129-175.

2. Varela, F.; Maturana, H., "Time course of excitation and inhibition in the vertebrate retina", en *Exp. Neurol,* 1970, 26: 53-59.

me guió en una lectura disciplinada de Teilhard de Chardin. En la búsqueda de una formación más sistemática, al cambiarme a la Facultad de Ciencias en 1966, me inscribí también en la Licenciatura en Filosofía en el antiguo Instituto Pedagógico de la calle Macul y comencé a participar regularmente en las lecturas guiadas por Roberto Torreti en el Centro de Estudios Humanísticos de la Escuela de Ingeniería. Las grandes polémicas ideológicas del Pedagógico no me interesaron tanto como lo que pude describir gracias a las clases de Francisco Soler que resonaban con la formación alemana de Torreti, y que se plasmaban en las colecciones de la biblioteca del Centro. Encontré así la fenomenología europea, y comencé una lectura, que se continúa hasta hoy, de Husserl, Heidegger, y Merleau-Ponty. Por primera vez me parecía encontrar en estos autores una preocupación por el tematizar la experiencia vivida que considero fundamental.

El segundo y fulgurante descubrimiento de estos años fue la naturaleza *social* de la ciencia. Le debo a Félix Schwartzman el haberme introducido temprano a todo ese mundo. En su curso de la Facultad de Ciencias conocí lo que hasta entonces era el saber de una minoría en Chile, los trabajos de la escuela francesa de historia y filosofía de las ciencias: Alexandre Koyré (sobre todo), Georges Canguilhem, y Gastón Bachelard. En todos estos autores se expresa la convicción no intuitiva de que las ideas científicas se hacen y cambian de manera abrupta y no por una suerte de acumulación de "evidencia puramente empírica", que se sustentan de imágenes e ideas que no son dadas e inmutables y que cada época es ciega a los fundamentos de lo que toma por cierto y evidente. El gran público se hizo consciente de todo esto gracias al famoso libro de Thomas Kuhn[3] que es imposible sin los antecedentes de la escuela francesa a la que Kuhn cita con reverencia. Para mis cortos 19 años, estas lecturas guiadas de Schwartzman sobre el quehacer científico me sacaron para siempre de la posición de ingenuo aprendiz a ser también un crítico de lo que yo recibía en mi formación profesional.

---

3 Kuhn, T., *The structure of scientific revolutions,* Harvard University Press, 1970.

La Facultad de Ciencias en aquella época pionera, hacía pocas concesiones en el nivel de formación matemática. En mi primer día de clases, sin decir una palabra, el profesor comenzó a escribir: "Sea E un espacio vectorial; los axiomas de E son:..." Después del primer shock para ponerse a nivel, descubrí en las matemáticas un lenguaje y una manera de pensar que me fascinaron. Al mismo tiempo descubrí gracias a Heinz von Foerster, el mundo de la cibernética, los modelos y la reflexión sistémica. Heinz es uno de los fundadores de todo ese universo de discurso y aunque no lo conocí en persona sino en 1968, se transformó en un personaje de gran importancia para mí. En sus p*apers,* que circulaban por el laboratorio de calle Independencia, aparecían títulos que me maravillaban, tales como *Historia natural de las redes neutrales* u otros más intraducibles como *A circuity of clues for platonic ideation*[4]. Encontré en estas ideas un instrumento para expresar las propiedades de los fenómenos biológicos, más allá de sus particularidades materiales.

Era una manera de pensar que había sólo aparecido en los años cincuenta, más claramente con la publicación del libro *Cybernetics* de Norbert Wiener (1962) y bajo la influencia de otro gran personaje de MIT, Warren Mc Culloch[5], a quien Humberto había conocido en 1959-60 cuando trabajaba en MIT. Wiener, McCulloch y von Foerster fueron pioneros de la *conjunción* de la reflexión epistemológica, la investigación experimental y la modelización matemática.

Entre otras cosas, aparecía aquí expresado de manera contagiosa que la metáfora del computador no era lo único que había a la mano para pensar en el operar del sistema nervioso.

4 Para una selección de estos y otros artículos véase: Von Foerster, H., *Observing systems: selected papers,* Interscience, California, 1979.

5. Una selección de sus trabajos más importantes apareció recién en 1975: Mc Culloch, W.S., *Embodiments of mind,* MIT Press.

Entrada en la epistemología experimental

El aprendizaje del oficio de neurobiólogo no era lo único que pasaba en el subterráneo de Independencia. Humberto había entrado francamente en un período de cuestionamiento de ciertas ideas dominantes en neurobiología, y la discusión, la lectura y el debate eran cotidianos, fomentados por la presencia de Gabriela Uribe, médico de clara inclinación epistemológica que trabajaba con Maturana en esa época. Eran tiempos de búsqueda y discusión para poner en foco lo que aparecía como una insatisfacción y una *anomalía.* La insatisfacción principal apuntaba al hecho de que la noción de información, supuestamente clave para entender del cerebro y el conocimiento, no aparecía jugando un papel explícito en el proceso biológico. Humberto intuía que los seres vivos son, como decía en aquella época, 'auto referidos', y que de alguna manera el sistema nervioso es capaz de generar sus propias condiciones de referencia. Se trataba de hacer una reformulación que condujera a una "epistemología experimental", un feliz término introducido por McCulloch. Gabriela y Humberto habían comenzado un estudio de ciertos efectos cromáticos similares a los descritos por E. Landen 1964, y que se transformaron en el tópico alrededor del cual se realizaba un primer intento de reformular la percepción visual como no-representacional.

El tiempo de mi formación en Chile terminaba. El Departamento de Biología de la Facultad de Ciencias me ofreció apoyo para obtener una beca de la Universidad de Harvard para hacer un doctorado (aunque sólo había completado 4 años de universidad, dos en Medicina y dos en la Facultad). Empecé a cerrar mi vida de estudiante en Chile, consciente de partir con un centro de interés claro en la epistemología experimental, y con sus tres pilares vivos en mi imaginación.

Harvard y la crisis del 68

Partí a Harvard un día 2 de enero de 1968, en un avión de Braniff, y leyendo un texto de Koyré sobre Platón. Llegué a Cambridge

con una gran tormenta de nieve, sin casa, lejos de hablar inglés fluidamente, y con la espada de Democles sobre mí: si no rendía en todo con 'A', la beca sería suspendida. Fueron unos primeros meses duros, pero una vez instalado, sabiendo ya moverme en este nuevo reino, me lancé de lleno a escuchar cursos y seminarios de todo tipo: en antropología (los estudios sobre la etología natural de primates comenzaba), en evolución (S. Gould acababa de llegar a Harvard y contrastaba con E. Mayr, el clásico), en matemáticas (la teoría de sistemas dinámicos no lineales se descubría en esa época), y en filosofía y lingüística (Chomsky era la figura dominante junto a Putnam y Quine). Encontré en Cambridge bibliotecas hasta entonces de fábula, bien provistas y abiertas a toda hora. Tenía la impresión de haber saltado de galaxia, y no recuerdo un solo día en que no sintiera las ganas de absorber como heliogábalo todo lo que tenía a la mano.

Muy luego me di cuenta, con gran sorpresa, de que, respecto a mis compañeros de generación en los estudios doctorales, mi visión de la ciencia y mis intereses eran francamente más heterodoxos y maduros que la mayoría. Es más, me di cuenta que ponerse a hablar con mis profesores de problemas epistemológicos como estaba habituado a hacerlo en Santiago, era mal visto. Lo mismo ocurrió cuando intenté buscar cultivar mis intereses en biología teórica. Lo que había sido la escuela de MIT, en 1968 había ya desaparecido, con Mc Culloch jubilado y sin nadie que tomara su lugar. Mi único punto de referencia continuaría siendo Von Foerster a quien visité varias veces en el Biological Computer Laboratory en la Universidad de Illinois en Urbana, un activo y productivo centro que él dirigía en esos años. Fue fácil concluir que mi búsqueda intelectual tendría que ser en dos tiempos: el oficial y el privado.

Oficialmente me hice alumno de Keith Porter, en cuyo laboratorio aprendí a trabajar en biología celular, y de Torsten Wiesel, quien poco después recibiría el Premio Nobel por sus trabajos sobre el "procesamiento de información" en la corteza visual. Orienté mi interés a aspectos comparativos de la visión y me puse a trabajar en la estructura funcional de los ojos de los insectos, que sería el tema

de mi Tesis. Para comienzos de 1970 había ya publicado cuatro artículos en el tema, y aprobé mi Tesis en abril de 1970.

Fuera del laboratorio y extraoficialmente, por primera vez me movía en un mundo enormemente más vasto que el de Santiago, con jóvenes de otra cultura, donde se mezclaban las nacionalidades y las razas. El azar quiso que esos años portaran los míticos eventos que marcaron a toda mi generación. Lo que comenzara en París la noche del 10 de mayo de 1968 correspondía con el M*ovement* norteamericano, nucleado por la oposición a la guerra de Vietnam. A los muertos en Kent State siguieron las primeras huelgas estudiantiles a las que me uní, con momentos dramáticos como la noche en que la policía nos sacó a palos de Harvard Yard. Los años en Cambridge fueron para mí el descubrimiento de mi inserción social ciudadana y de la posibilidad de hacerme responsable de cambios en mi entorno social. Fue un reencuentro, a la distancia, con mis raíces en América Latina a través de mis amigos del *Movement* que exaltaban la revolución cubana. No sólo era la ciencia lo que me ocupaba, era también el sueño de pensar en una América Latina nueva, propia de nuestra generación.

El haberme descubierto como animal social y político, acentuó la necesidad de guardar silencio en público sobre mis verdaderos intereses en los círculos oficiales. Fiel a la idea de ciencia como actividad que se hace y se crea a saltos y por innovaciones atrevidas, cultivé con mis camaradas de generación la intención de volver a Chile a hacer una ciencia distinta, donde las anomalías que ya había presentido en Chile y que se acentuaban en EE.UU., pudieran ser transformadas en práctica científica. Hacer ciencia original y propia parecía lo mismo que el compromiso con mi historia y mis orígenes.

Me gradué de doctor en Biología en junio de 1970. Contra el reclamo de mis profesores, decliné un cargo como investigador de Harvard y otro como Profesor Asistente en otra universidad americana. Decidí aceptar el cargo que me ofrecía la Facultad de Ciencias, justamente deseosa de recuperar el esfuerzo que se había puesto en mi formación. Volví a Chile el 2 de septiembre de 1970,

y la elección de Allende dos días más tarde me pareció mi segunda y verdadera graduación. Por fin el trabajo podía comenzar en pleno, con problemas claves bien delimitados, con la seguridad de ser tan preparado y competente como el que más en la escena científica mundial, y con el contexto de trabajar en una inserción en la que había un futuro por construir. Esta convergencia de circunstancias fue absolutamente decisiva. Es con todos estos ingredientes de la situación a septiembre de 1970 que puedo ahora volver a la especificidad de la noción de auto*poiesis* y su gestación.

LA GESTACIÓN DE LA IDEA

Cerniendo el problema

El antecedente directo de la gestación de la *autopoiesis* es el texto de Maturana escrito hacia mediados de 1969 originalmente titulado *Neurophysiology of cognition.* Humberto había continuado su propio camino de interrogación sobre lo inadecuado de la idea de información y representación para entender el sistema biológico. Pasó a visitarme en varias ocasiones a Cambridge y, como en Santiago, tuvimos largas conversaciones. En el semestre de primavera de 1969, Heinz von Foerster lo invitó por algunos meses al Biological Computer Laboratory, ocasión que coincidió con una reunión internacional de la Wenner Green Foundation bajo el tema "Cognition: A multiple view", un titulo visionario a la luz del desarrollo enorme de las hoy llamadas ciencias cognitivas, pero hasta entonces no formuladas como campo de investigación científico.

Humberto preparó para esa reunión el texto mencionado, dando por primera vez una expresión clara y atractiva, a sus ideas maduradas hasta entonces, para poner más en claro lo que hasta entonces aludía como el carácter autorreferido de los seres vivos, y para definitivamente identificar la noción de representación como el pivote epistemológico que había que cambiar. En su lugar era necesario poner al centro de atención la concatenación interna de los procesos neuronales, y describir al sistema nervioso como un sistema "cerrado" como dice el texto. Este artículo marca un salto

importante, y todavía hoy creo que inicia de manera indiscutible un giro en una nueva dirección. Recuerdo haber ido a visitar a Humberto a Illinois y haber discutido varias partes difíciles del texto mientras el parto se concluía. El texto apareció poco después[6], y el artículo se abre con un párrafo de agradecimiento a las muchas conversaciones con Heinz y conmigo. Poco después Humberto retrabajó este texto a uno más definitivo que pasó a llamarse *Biology of cognition.*

En este texto se toca sumariamente una idea que me venía intrigando desde antes, y que como ayudante del curso de biología celular que enseñaban George Wald y James Watson en Harvard me había aparecido como no muy claro, como una anomalía: se hablaba de la constitución molecular de la célula, y se usaban términos como automantención, pero nadie, ni aun los dos premios nobeles reunidos, sabían lo que se quería decir con ello. Lo que es peor, cuando empujaba yo la discusión a la hora del almuerzo, la reacción habitual era un típico "Francisco, siempre metiéndose en filosofía". Mis notas de aquella época incluyen varios intentos de cernir la autonomía básica del proceso celular como base de la autonomía de lo viviente. Hacia fines de 1969 apareció en el escaparate de Shoenhoff's Foreing Books en Cambridge, el op*us magnum* de Jean Pieaget titulado *Biologie et connaissance*[7], donde se apunta claramente a una necesidad de revisar la biología en la dirección de "l'autonomie du vivant", pero el lenguaje de Piaget y sus idiosincrasias me dejaban insatisfecho.

En su artículo, Humberto hacía el vínculo entre el carácter circular de los procesos neuronales y el hecho que el organismo es también un proceso circular de cambios metabólicos, como se ilustraba con referencia a un artículo reciente de Commoner aparecido en *Science* que discutía los nuevos avances de la bioquímica del metabolismo y su evolución. La pregunta que se cernía entonces era ésta: ¿si dejamos de lado por el momento la organización del

6. Garvin, P. (ed.), *Cognition: a multiple view,* Spartan Books, Washington, 1970.

7. Piaget, J., *Biologie et connaissance,* Gallimard, París, 1969; traducción española apareció en Siglo XXI, Buenos Aires en 1972.

sistema nervioso y centramos la pregunta sobre la autonomía de lo vivo en su forma celular, qué podemos decir? Esta reflexión sobre la naturaleza circular del metabolismo en los seres vivos y su relación con el operar cognitivo, aunque ocupa una corta página en la versión definitiva de *Biology of cognition,* sería un punto focal desde donde arrancaría el desarrollo de la noción de *autopoiesis.*

Estos eran los meses finales de 1970. Estaba yo de vuelta en Chile, ya que el Departamento de Biología me había pedido que me hiciera cargo del curso introductorio de Biología Celular para nuestros nuevos alumnos. Con Maturana éramos ahora colegas en el Departamento de Biología, vecinos de oficina en las barracas "transitorias" (pero todavía utilizadas) del nuevo campus de la Facultad de Ciencias en la calle Las Palmeras, en Macul. Estaba todo en su lugar para lanzarse a explorar la pregunta sobre la naturaleza de la organización mínima de lo vivo, y no perdimos tiempo. En mis notas los primeros esbozos más maduros aparecen ya a fines de 1970, y hacia fines de abril de 1971 aparecen más detalles junto con un modelo mínimo que sería más tarde simulado en el computador. En mayo de 1971 el término *autopoiesis* figura en mis notas como resultado de la inspiración de nuestro amigo José M. Bulnes, quien acababa de publicar una tesis sobre el Quijote donde se utilizaba la distinción entre praxis y poesis. Una nueva palabra nos convenía porque queríamos designar algo nuevo. Pero la palabra sólo adquirió poder asociada al contenido que nuestro texto le asigna; su resonancia va más allá del mero encanto de un neologismo.

Fueron meses de discusión y trabajo casi permanentes. Algunas ideas las probaba con mis estudiantes del curso de Biología Celular, otras con colegas en Chile. Era claro que nos embarcábamos en una tarea que era conscientemente revolucionaria y antiortodoxa, y que ese coraje tenía todo que ver con el estado de ánimo de Chile donde las posibilidades se abrían a una creatividad colectiva. Los meses que llevaron a la configuración de la *autopoiesis* no son separables del Chile de entonces.

En el invierno de 1971 sabíamos que teníamos un 'concepto importante' entre las manos y decidimos escribirlo. Un amigo nos prestó su casa en la playa de Cachagua. Fuimos en dos ocasiones entre junio y diciembre. Los días en la playa se dividían en largas caminatas, y sobre todo un ritmo monástico de escritura que normalmente iniciaba Humberto y retomaba yo. Al mismo tiempo iniciaba yo una primera versión (que revisaba Humberto) de un artículo más breve que expondría las ideas principales con la ayuda de la simulación de un modelo mínimo (que llamamos "Protobio", como detallo más adelante). Hacia el 15 de diciembre (otra vez según mis notas de 1971), teníamos una versión completa de un texto en inglés llamado: *Autopoiesis: the organization of living systems.* En su versión dactilografiada eran 76 páginas, de las cuales se hicieron algunas docenas de copias en el antiguo método de roneo en tinta azul. Aunque hubo algunas modificaciones ulteriores, ese texto es lo que el lector tiene entre sus manos en su traducción española.

Como ha ocurrido a menudo en la historia de la ciencia, la dinámica creativa entre Maturana y yo fue una resonancia en espiral ascendente, en la que participaba un interlocutor ya maduro que aportaba su bagaje de experiencia y pensamiento previo, y un joven científico que aportaba ideas y perspectivas frescas. Como es evidente dadas las circunstancias, las ideas no surgieron en una conversación ni en dos, ni era una simple cuestión de hacer explícito lo que estaba ya dicho antes. Lo que estaba en el trasfondo debió ser configurado en un salto *cualitativo.* Tales transiciones no son nunca simples, ni es posible retratar cómo ocurrieron de manera exacta, porque es siempre una mezcla de pasado y presente, de talentos y debilidades, de imaginación e inspiración. La noción madura de *autopoiesis* tenía, como hemos visto, claros antecedentes, pero entre los antecedentes y una idea madura hay un salto que es crucial. Y así como Franklin no es la doble hélice de Watson & Crick, ni Lorentz es la relatividad especial, los antecedentes claves de la *autopoiesis* no son reducibles a la expresión madura de la idea, como es fácil ver comparando los textos publicados. Es un ejemplo límpido de lo que había ya aprendido con mis maestros franceses: que la ciencia tiene

discontinuidades, que no funciona por acumulación empírica progresiva, y que es inseparable de su contexto histórico social.

Dejemos por el momento la filigrana histórica e identifiquemos cuál es, más precisamente, la especificidad del concepto y cómo se constituye como un salto cualitativo.

La especificidad de la autopoiesis

¿Qué era lo que habíamos logrado en esos meses de intensa producción? ¿Por qué el concepto habría de tener una resonancia más allá de las barracas de la Facultad de Ciencias? Para poder responder, necesito la paciencia del lector para poder decir qué es la *autopoiesis* y diferenciarla de lo que no es. Esto por dos razones. La primera, porque lo que esté dicho en el texto de este libro ha tenido una maduración en las dos décadas que lo siguieron que lo aclaran y lo hacen más terso que cuando fuera escrito por primera vez. Segundo, porque es sobre la base de mi comprensión actual que hablo aquí sobre su historia y su devenir.

Lo que demarca el trabajo hecho en este texto es que por primera vez se articulan explícitamente las ideas siguientes:

0. El problema de la *autonomía* de lo vivo es central y hay que cernirlo en su forma mínima, en la caracterización de la unidad viviente.

1. La caracterización de la unidad viva mínima no puede hacerse solamente sobre la base de componentes materiales. La descripción de la *organización* de lo vivo como configuración o *pattern* es igualmente esencial.

2. La organización de lo vivo es, en lo fundamental, un mecanismo de constitución de su *identidad* como entidad material.

3. El proceso de constitución de identidad es *circular:* una red de producciones metabólicas que, entre otras cosas, producen una membrana que hace posible la existencia misma de la

red. Esta circularidad fundamental es por lo tanto una au*toproducción* única de la unidad viviente a nivel celular. El término *autopoiesis* designa esta organización mínima de lo vivo.

4. Toda interacción de la identidad autopoiética ocurre, no sólo en términos de su estructura físico-química, sino que también en tanto unidad organizada, esto es, en referencia a su identidad autoproducida. Aparece de manera explícita un punto de *referencia* en las interacciones y por tanto la emergencia de un nuevo nivel de fenómenos: la constitución de significados. Los sistemas autopoiéticos inauguran en la naturaleza el fenómeno *interpretativo.*

5. La identidad autopoiética hace posible la evolución a través de series reproductivas con variación estructural con *conservación* de la identidad. La constitución identitaria de un individuo precede, empírica y lógicamente, el proceso de *evolución.*

Estos cinco puntos entrelazados expresan la especificidad de la *autopoiesis* como noción, su ruptura con concepciones anteriores y, en mi opinión, el fundamento de por qué ha encontrado resonancia en estas dos décadas. En efecto, la idea condensa de una manera compacta y casi holográfica tres conceptos que están al centro de las preocupaciones de varias disciplinas científicas actuales: la neurobiología y la biología evolutiva, las ciencias cognitivas y la inteligencia artificial, las ciencias sociales y de la comunicación:

Hay en la naturaleza propiedades radicalmente *emergentes,* que surgen de sus componentes de base, pero que no se reducen a ellos. La vida celular es un caso ejemplar de una tal propiedad emergente, y sobre esta base puede definirse *lo vivo* de una manera precisa y aún formalizable.

Toda serie evolutiva es secundaria a la *individuación* de los miembros de la serie. El proceso de individuación contiene capacidades emergentes o internas que hacen que la serie evolutiva no se explica sólo sobre la base de una selección externa, sino requiere también de las propiedades intrínsecas de la autonomía de los individuos que la constituyen.

El fenómeno *interpretativo* es una clave central de todos los fenómenos cognitivos naturales, incluyendo la vida social. La significación surge en referencia a una identidad bien definida, y no se explica por una *captación* de información a partir de una exterioridad.

Lo que esta idea *no* evoca es la deriva histórica de sistemas celulares terrestres, tal como los conocemos hoy en su expresión mínima bacteriana. En particular la *autopoiesis* establece las condiciones necesarias para una serie evolutiva porque determina un tipo de individuos, pero no se pronuncia sobre la manera como esos individuos adquieren cambios estructurales que les permiten una deriva evolutiva rica y diversa. Es aquí donde entran a jugar un rol clave los ácidos nucleicos (ADN, ARN) y las proteínas (las llamadas moléculas con *información)* en tanto soporte de la herencia celular, lo que probablemente comenzó con el *mundo del ARN*. La discusión sobre el origen de la vida habitualmente se centra sobre esta serie de cambios estructurales[8]. En contraste, l*a autopoiesis* sólo pretende establecer la clase de individuos con la que esa serie evolutiva comienza y se origina: se trata del criterio de demarcación entre los primeros seres vivos y la sopa primordial que los precede, ni más ni menos. Mantener estas distinciones a la vista permite evitar muchas discusiones estériles.

Una idea y dos textos

Lo que acabo de perfilar no era evidente, insisto, en 1971. Como es inevitable, la comprensión se desenvuelve a lo largo del tiempo y en la medida de sus efectos. No sorprende entonces el que el texto que concluimos hacia fines de 1971 no tuvo una aceptación inmediata. De hecho fue enviado al menos a cinco editores y revistas, y todos sin excepción lo consideraron impublicable. Recuerdo que en enero de 1972, mi ex profesor Porter me invitó de visita al nuevo Departamento de Biología de la Universidad de Boulder, donde

8. Sobre estos trabajos véase Deamer, D.; Fleschaker, G. (Eds.), *The origins of life: the central concepts*, Jonathan Cape, Boston, 1994.

con entusiasmo di una charla titulada: "Cells as autopoietic machines". La recepción fue fría y distante, como lo fue también la de colegas de Berkeley que visité por la misma época.

Las dificultades de publicación, añadidas al momento político por el que pasaba Chile hacia fines de 1972, me hacían sentirme alienado del mundo científico internacional. Por lo mismo, la ocasional recepción entusiasta de ciertas personas a quienes yo respetaba fue de un enorme valor. El primero en tener una percepción clara de las posibilidades de la idea fue naturalmente nuestro amigo Heinz en EE.UU., con quien había comunicación constante y quien vino a Chile durante esos años. Otro cibernético y sistemista ya célebre que tuvo una reacción positiva fue Stafford Beer que venía regularmente a Chile. En efecto, Fernando Flores lo había contratado a nombre del gobierno para echar a andar un sistema revolucionario de comunicación y regulación de la economía chilena inspirado en el sistema nervioso, que pasó a llamarse Proyecto Cinco. Beer respondió con tal entusiasmo a lo que el texto planteaba que decidimos pedirle un Prefacio que él accedió a escribir inmediatamente. En enero de 1972, con una copia todavía fresca del manuscrito fui invitado a México por Iván Illich a su centro CIDOC en Cuernavaca. El día de la llegada, le di el manuscrito y, a la mañana siguiente, me quedó grabada su reacción: "Es un texto clásico. Ustedes han logrado poner la autonomía al centro de la ciencia". A través de Illich, el texto llegó a manos del famoso psicólogo Erich Fromm, quien me invitó a su casa-retiro a discutir del nuevo concepto, que él incorporó de inmediato al libro que escribía por entonces[9]. En Chile mismo, Fernando Flores y otros colegas del Proyecto Cinco fueron también un público atento a nuestra manera de pensar. Trabamos con Flores lo que habría de ser una fructífera amistad, y muchos años más tarde la *autopoiesis* figuraría entre los conceptos importantes que él utilizaría para desarrollar sus propias ideas. Es difícil imaginar todo lo que significó para mí en esa época el encontrar receptividad en personas de esta calidad.

---

9. Se trata de su libro The *anatomy of agression.*

Entretanto el texto seguía haciéndose rechazar por una lista creciente de editores extranjeros. Era natural entonces dirigirse a la Editorial de nuestra Universidad, y a fines de 1972 firmamos un contrato que incluía la traducción del texto por doña Carmen Cienfuegos. *De máquinas y seres vivos: una teoría sobre la organización biológica* se imprimió en abril de 1973. El texto original en inglés no aparecería sino hasta 1980, cuando la idea había ya adquirido una cierta popularidad en la prestigiosa serie "Boston studies on the philosophy of science", con una Introducción firmada por Maturana, el texto *Biology of cognition,* el prefacio de Beer, y el texto en cuestión *Autopoiesis: the organization of living systems*[10]. Según me informa el editor, este libro ha sido el *best-seller* de la colección.

El destino del breve artículo escrito en paralelo a este texto sufrió una historia similar. Como mencioné antes, además de una presentación sucinta de la noción de *autopoiesis* la intención del artículo era de ayudar a la claridad expositiva a través de un caso mínimo de *autopoiesis.* Ya hacia fines de 1970 habíamos llegado a la conclusión de que un caso simple de auto producción requeriría dos reacciones: una de polimerización de elementos de membrana, la otra de generación "metabólica" de monómeros. Esta última debía ser una reacción catalizada por un tercer elemento preexistente en la sopa de reacción. Concebido este esquema de reacción, parecía evidente probar una simulación de este caso mínimo (que pronto pasó a llamarse *Protobio* en nuestra conversación) utilizando autómatas celulares (o de *teselación,* como se decía entonces), un útil de modelización introducida en los años 50 especialmente por John von Neuman. Con la colaboración de Ricardo Uribe de la Escuela de Ingeniería, las simulaciones dieron rápidamente los resultados que la intuición nos hacía esperar: la emergencia espontánea en este mundo bidimensional artificial de unidades que se auto distinguían a través de la formación de una membrana, y que mostraban una capacidad de auto reparación. El *paper* fue enviado a varias revistas incluyendo *Science y Nature,* con resultados

10. Maturana, H.; Varela, F.J., *Autopoiesis and cognition: the realization of the living,* BSPS, vol. 42, D. Reidel, Boston, 1980.

semejantes al texto del libro: rechazo completo. Heinz vino a Chile de visita en el invierno de 1973, y nos ayudó a reescribir el texto de manera significativa. Se lo llevó a EE.UU. bajo el brazo y lo envió al editor de la revista *Biosystems* de la cual era miembro del comité editor. El *paper* sufrió algunos comentarios duros de los revisores, pero poco después fue aceptado y finalmente apareció a mediados de 1974[11]. Este artículo es importante de mencionar aquí porque fue la primera publicación de la idea de *autopoiesis* en inglés en el mundo internacional, lo que llevó a la comunidad internacional a hacerse cargo de la idea, y porque anticipó en 20 años lo que habría de llegar a ser el explosivo campo hoy llamado de la vida artificial y los autómatas celulares, como explico más adelante.

La visita de Heinz en julio de 1973 tuvo lugar en el medio de la tormenta que se avecinaba y que nos sumía a todos en una atmósfera de crisis permanente con desesperados intentos por estabilizar un país que se dividía en dos. Como militante comprometido con el gobierno del Presidente Allende, a partir del 11 de septiembre me encontré amenazado. Inteligencia militar vino a la Facultad con listas de ex partidarios, y en dos ocasiones patrullas nocturnas vinieron a buscarme a mi domicilio donde ya no acudía a dormir. Fui exonerado de mi cargo universitario por órdenes "superiores". Con mi familia decidimos vender todo y partir. La mayor parte de mis colegas de la Facultad de Ciencias se dispersaban también por el mundo. Con la diáspora de los científicos de la Facultad, se acababa una época de la ciencia en Chile, una etapa importante de mi vida personal, y con ella el contexto que dio origen a la idea de *autopoiesis.* Pero naturalmente la idea habría de tener nuevos avatares, sobre todo fuera de Chile.

11. Varela, F.; Maturana, H.; Uribe, R., "Autopoiesis: the organization of living systems, its characterization and a model", en *Biosystems,* 1974, 5: 187-196.

Devenir de la autopoiesis en sentido estricto

¿Cuál fue el devenir de la idea de *autopoiesis al* interior de las ciencias? Poco después de la aparición del artículo en *Biosystems* en 1974, la idea empezó a hacer su camino con cierta inercia en los medios científicos. No viene al caso hacer aquí una historia detallada, quiero sólo dar de modo indicativo algunos hitos.

Aquellos que se ocupaban de teoría de sistemas fueron los primeros en reaccionar, y ya en 1976 en la reunión internacional de sistémica en New York había una sesión especial llamada *Autopoiesis* (en la que fui conferenciante invitado). El primer libro dedicado exclusivamente a la idea apareció poco después[12]. Habrían de haber muchos otros libros y artículos en años posteriores. Más importante para mí, fue que la idea empezó a suscitar interés entre los biólogos, muy especialmente a través de Lynn Margulis, destacada investigadora del origen de la vida y la evolución celular. A partir de comienzos de los años 80, Margulis adopta la *autopoiesis* como el criterio para definir el origen de los seres vivos y difundió la idea a muchos otros científicos activos en el área. Bajo Margulis, Gail Heischaker escogió el tema para su tesis doctoral y publicaría más tarde una serie de artículos originales sobre el tema.

Aunque a partir de entonces, a nivel teórico y biológico la idea de *autopoiesis* se encarnaba en el discurso científico, por muchos años tenía yo la esperanza de que pudiera servir de guía para la síntesis de sistemas precelulares. Estos desarrollos habrían de esperar los años 90 cuando Luigi Luisi y su grupo en Zurich decidieron lanzarse al desafío de la construcción de sistemas autopoiéticos sintéticos sobre la base de micellas lipídicas, haciendo

12. Zeleny, M. (ed.), *Autopoiesis: a theory of the living organization, N.* Holland, Nueva York. Véase también Benseler, F.; Hejl, P.; Kock, W., *Autopoietic systems,* Campus Verlag, Frankfurt, 1980, y los números especiales dedicados a la *autopoiesis* de: *Irish J. Psychology,* 1988, *Cybernetics* 1985, *Int. J. Gen systems, 1992.*

uso de un *know-how* experimental considerable[13]. Los resultados muestran claramente que la circularidad autopoiética puede implementarse en un sistema químico real, sometido a restricciones similares a las de los primeros sistemas celulares. Como señalaba el editorial de *Nature,* comentando un artículo de Luisi que aparecía en el mismo número, la síntesis de sistemas autopoiéticos artificiales representa 'el haber completado una etapa más para resolver el misterio del origen de la vida"[14].

Estos últimos años han sido también testigos de una nueva corriente de investigación interdisciplinaria llamada *Vida artificial,* continuación natural de la cibernética de los años 50, uno de cuyos objetivos es la simulación y realización de sistemas vivos a varios niveles, desde el celular hasta el robótico[15]. Uno de los postulados más repetidos de esta corriente, hoy en día muy publicitada en medios periodísticos, es que es una organización lo que permite definir la vida, y no los componentes, por muy sofisticadas que sean sus propiedades enzimáticas o replicativas. Esa es, por cierto, una intuición que guiaba nuestra búsqueda en 1971. Más aún, los autómatas celulares que usáramos en el protobio, se transformaron en las manos de la *vida artificial,* en la herramienta predilecta de simulación de toda clase de propiedades biológicas.

La autopoiesis como metonimia

No puedo omitir aquí un comentario sobre otra dimensión de la expansión de la idea de *autopoiesis* más allá de la biología hacia las ciencias humanas, donde ha suscitado un interés inusitado. Pienso

---

13. Luisi, L.; Varela, F., "Self-replicating micelles as a chemical version of a minimal autopoietic system" en *Origins of Life,* 1989, 19: 633-643; Bachman et al., "Autocatalytic self-replicating micelles and models for prebiotic structures", en *Nature,* 1992, 357: 57.

14. *Nature* 354: 351, 1991.

15. Véase Langton, C. (ed.), *Artificial life I, Addison Wesley,* Nueva Jersey; Varela, F.; Bourgine, P. (eds.), *The practice of autonomous systems: the first European conference on artificial life,* MIT Press/Bradford Books, Cambridge, 1992; para una introducción en español, véase Fernández, J.; Moreno, A., *La vida artificial,* Editorial Eudema, Madrid, 1992.

que en estos casos la *autopoiesis* aparece jugando un rol metafórico, o más precisamente, *metonímico* Esta tendencia ya se planteaba en el prefacio que Stafford Beer escribiera en 1972, donde afirma que es "evidente" que la idea puede extenderse para caracterizar un sistema social. Ya en esa época tenía yo una posición escéptica al respecto, como lo señala el mismo Beer.

En los años que siguieron, este uso metonímico tomó fuerza en dominios tan diversos como la sociología, en los escritos del famoso sociólogo alemán Niklas Luhman[16], la teoría jurídica[17], la teoría literaria[18], así como una extensa literatura en el campo de la terapia familiar sistémica[19]. Toda esta profusión de interés ha sido para mí fuente de sorpresa. Después de años de escuchar los argumentos y los usos de la idea en varios de estos campos, he llegado a algunas conclusiones generales de las que quiero dejar constancia brevemente.

Quiero distinguir en esta literatura secundaria dos modos de transposición de la idea original: (1) una utilización literal o estricta de la idea, (2) una utilización por continuidad. Con el primer modo me refiero al hecho que ha habido intentos repetidos de caracterizar, por ejemplo, una familia como un sistema autopoiético, de manera que la noción se aplique en este caso estrictamente. Estos intentos se fundan, en mi opinión, en un abuso de lenguaje. En la idea de *autopoiesis* las nociones de red de producciones y de frontera tiene un sentido más o menos preciso. Cuando la idea de una red de procesos se transforma en "interacciones entre personas", y la membrana celular se transforma en el "borde" de un grupo humano, se incurre en un uso abusivo, como lo expresé en un comentario crítico que he publicado al respecto[20].

16. Kuhman, N., *Soziale systeme,* Suhrkamp, Frankfurt, 1984.

17. Teubner, G., *Law as an autopoietic system,* Blackwell, Oxford, 1993.

18. Paulson, W., *The noise as culture,* Cornell Univ. Press, Itaca, 1988.

19. Véase como un ejemplo Elkaim, M., *Si tu ne m'aimes pas,* Seuil, París, 1992.

20. Varela, F., "On the circulation of concepts from a biology of cognition and systemic family therapy", en *Fam. Process,* 28: 15-24, 1989.

El uso de la *autopoiesis* por continuidad es otro: se trata de tomar en serio el hecho de que la *autopoiesis* busca poner la autonomía del ser vivo al centro de la caracterización de la biología, y abre al mismo tiempo la posibilidad de considerar los seres vivos como dotados de capacidades interpretativas desde su mismo origen. Es decir permite ver que el fenómeno interpretativo es *continuo* desde el origen hasta su manifestación humana. En lo fundamental, yo estoy de acuerdo con este uso y esta extensión posible. En el panorama de ideas actuales es quizá una de las facetas más originales de este trabajo. Sin embargo, pienso que darle una argumentación y una expresión rigurosa a esta articulación requiere trabajo serio. Entre los ejemplos que me parecen convincentes destacan el que hace J.P. Dupuy en su análisis de los sistemas sociales[21], Winnograd & Flores en su análisis sobre la comunicación[22], y el texto más evocativo de W.I. Thompson[23]. Por desgracia, a menudo la idea de *autopoiesis* se la cita en este tipo de literatura de manera mucho más superficial.

En breve, creo que queda claro para el lector que, en lo fundamental, tengo un gran escepticismo sobre la extensión del concepto más allá del área para la que fue pensado, es decir la caracterización de organización de los sistemas vivos en su expresión mínima. Aunque no hay una razón a *priori,* después de todos estos años mi conclusión es que una extensión a niveles "superiores" no es fructífera y que debe ser dejada de lado, aun para caracterizar un organismo multicelular[24]. Por el contrario, el ligar la *autopoiesis* como una opción epistemológica más allá de la vida celular, al operar del sistema nervioso y los fundamentos de la comunicación humana, es claramente fructífero[25].

---

21. Dupuy, J.P., *Ordres et désordres,* Seuil, 1989.

22. Winnograd, T.; Flores, F., *Understanding computers and cognition,* Addison Wesley, Nueva Jersey, 1987.

23. Thompson, W.I., *Imaginary landscapes,* St. Martin Press, Nueva York, 1989.

24. Véase al respecto Varela, F.; Frenk, S., "The organ of form: towards a biological theory of shape", en *J. Soc. Biol. Struct.,* 1987, 10: 73-83.

25. Varela, F., "Organism: A meshwork of selfless selves", en Tauber, F. (ed.), *Organism and the origin of self,* Kluwer Assoc., Dordrecht, 1991, Pp. 79-107.

MÁS ALLÁ DE LA AUTOPOIESIS

Quiero terminar con un breve comentario sobre lo que ha sido mi camino intelectual después de 1973, independientemente de los avatares tempranos o tardíos, de la *autopoiesis*. Lo hago sobre todo porque quisiera aprovechar este Prefacio para dejar en claro que a mis ojos la *autopoiesis* fue una etapa importante y útil, pero tan sólo una pieza del puzzle más grande para entender la biología del conocer de una manera nueva. Ciertamente mi manera de pensar hoy no queda caracterizada como "la teoría autopoiética", contrariamente a lo que a veces se dice; es otro uso metonímico que sería mejor evitar.

Otra motivación para esbozar aquí estos hitos de mi propio desarrollo científico, es que después de los años de trabajo conjunto con Maturana en 1970-73, y un breve interludio de nuevas colaboraciones en mi segundo intento de vuelta a Chile en 1980-83, nuestros caminos intelectuales han divergido de múltiples maneras tanto en contenido, como en enfoque, como en estilo. Esto no debe sorprender; somos dos individuos diferentes y viviendo casi todo el tiempo en lugares distintos. Es importante, entonces, aclarar que nuestra colaboración en la creación de la *autopoiesis* no implica el que compartamos hoy una visión común en una supuesta "teoría autopoiética" unificada. Por cierto hay algún terreno compartido y que hemos expresado en un libro destinado al gran público, *El árbol del conocimiento*[26], que redactamos entre 1982-83 y que ha tenido un éxito significativo en los doce idiomas a que ha sido traducido. Por el contrario, el *corpus* de investigaciones que menciono a continuación representa trabajo científico original, y cuya responsabilidad me incumbe.

AUTORREFERENCIA Y CIERRE

La *autopoiesis* reposa sobre una concepción circular y autorreferencial de los procesos. Sin embargo por muchos años la autorreferencia recibió una atención marginal y más bien negativa. Una de las

---

26. Maturana, H.; Varela, F., *El árbol del conocimiento: las bases biológicas del entendimiento humano,* Editorial Universitaria, Santiago, 1984.

cosas que me han ocupado por períodos desde 1974 (y sobre todo trabajando con mi amigo y colega el matemático Jorge Soto-Andrade), es la de esclarecer la noción de autorreferencia como concepto formal y lógico bien fundado[27].

Junto con la búsqueda de un fundamento más claro a la circularidad, se me hizo claro también que la noción de *autopoiesis* es un caso particular de una clase o familia de organizaciones con características propias. Lo que tiene en común es que todas ellas dan al sistema en cuestión una dimensión autónoma. La base de esta conclusión fue sobre todo el repensar desde este nuevo ángulo las dos redes biológicas cognitivas más evidentes: el sistema nervioso (expresado en mi trabajo experimental de varias maneras) y el sistema inmunitario (que comenzó a ocuparme desde 1976). La idea es simple: sólo una circularidad del tipo de la *autopoiesis* puede ser la base de una organización autónoma. La caracterización de la clase de organización pertinente es lo que comencé a llamar el principio de cierre *operacional.* La palabra *cierre* la uso aquí en su sentido de operación al interior de un espacio de transformaciones, como es habitual en matemáticas, y no, por cierto, como sinónimo de cerrazón o ausencia de interacción, lo que sería absurdo. Lo que interesa es caracterizar una nueva forma de interacción mediado por la autonomía del sistema. Todas estas observaciones y conclusiones las resumí in *extenso* en un libro titulado *Principios de autonomía biológica* que apareció en 1979[28]. Ese libro continúa siendo la síntesis más completa del desarrollo de los fundamentos, las aplicaciones y los intentos de formalización de la noción de autonomía[29].

---

27. Varela, F., "A calculus for self-reference", en *Int. J. gen. sytems,* 1975, 2: 5-24; Varela, F.; Coguen, J., "The aritmethic of closure", en *J. Cybernetics,* 1978, 8: 291-324; Varela, F., "The extended calculus of indications interpreted as a three-valued logic", en *Notre Dame J. Formal Logic,* 1979, 20:141-146; Kaufman, L.; Varela, F., "Form dynamics", en *J. Soc. Biol. Struct.,* 1980, 3: 171-206; Soto Andrade, J.; Varela, F. ,"Self-reference and fixed points", en *Acta Applic. Matem.,* 1984, 2: 1-19.

28. Varela, F., *Principles of biological autonomy,* North-Holland, Nueva York, 1979.

29. Algunas de las formalizaciones algebraicas que propongo allí no me parecen hoy día tan útiles como entonces. Pero fueron un paso en la dirección adecuada, como lo muestran hoy los recientes trabajos de Walter Fontana en el Santa Fe Institute que retoma la noción de cierre y punto fijo en el contexto del cálculo de Church.

IDENTIDAD SOMÁTICA Y SISTEMA INMUNITARIO

A un nivel menos programático y más aplicado, una dirección de trabajo que ha sido de enorme fertilidad es un extensa reformulación del sistema inmunitario bajo los principios de organización autónoma. Este tema se abrió para mí al conocer en 1976 a Nelson Vaz en EE.UU.[30]. Sin embargo, no fue sino hasta mi instalación en París en 1986 y mi colaboración intensa con Antonio Coutinho del Instituto Pasteur, que esas instituciones han dado sus frutos. En pocas palabras, la idea central es la siguiente. La inmunología tradicional tiene como metáfora central la defensa contra los antígenos externos invasores. Esta es una transposición isomórfica de las ideas dominantes representacionistas en el sistema nervioso: los antígenos juegan el rol de *inputs* conteniendo información, la respuesta inmunitaria juega el rol de *output.* Nuestra proposición es en cambio: el sistema inmunitario es sobre todo un cierre operacional propio a los linfocitos y las regiones V-variables de las inmunoglobulinas, que permiten una identidad somática al organismo multicelular. Sólo secundariamente esta red desarrolla en el curso de la evolución capacidades defensivas tipo respuesta inmunitaria a infecciones masivas. Pero el corazón del funcionamiento del sistema es *constitutivo* de la identidad somática mediante la provisión de una red de inter comunicaciones al interior del paisaje celular y molecular del organismo, y no una serie de respuestas de anticuerpos dirigida al exterior. Se trata, por decirlo brevemente, de un verdadero sistema cognitivo del cuerpo.

Estas ideas las hemos expresado en numerosos trabajos[31], y han tenido un impacto en el mundo de la inmunología tradicional.

---

30. Vaz, N.; Varela, F., "Self and non-sense: an organism centered approach to inmunology", en *Medical Hypothesis,* 1978, 4: 231-267.

31. Véase en particular Varela, F.; Coutinho, A; Dupire, B.; Vaz N., "Cognitive networks: immune, neural, and otherwise", en Perelson, A. (ed.), *Theoretical Immunology, Part 1* (SF1 Series on the Science of Complexity), Addison-Wesley, Nueva Jersey, 1988, pp. 359-375; Stewart, J.; Varela, F., "Exploring the connectivity of the immune network", en *Immunol, Reviews n'* 110, 1989, pp. 37-61; Varela, F.; Coutinho, A., "Second generation immune networks", en *Immunol Today,* 1991, 12: 159-167; Stewart, J.; Varela, F., "Morphogenesis in shape space: elementary metadynamics of immune networks", en *J. Theoret. Biol.,* 1991, 153: 477-498; Bersini, H.; Varela, F., "Learning and the immune network; Reinforcement, recruitment and their applications", en Patron, G. (ed.), 1993, *Biologically Inspired Computation,* Chapman and Mill, Londres (in press).

Más interesante tal vez es que el paso entre la revisión del fenómeno inmunitario y su expresión en resultados experimentales nuevos e inesperados ha sido muy rápido. Así por ejemplo, nuestro enfoque llevó a poner en evidencia por primera vez que el sistema como un todo tiene una rica dinámica temporal en la composición de inmunoglobulinas que expresa, entre otras cosas, la diferencia entre normalidad y enfermedades autoinmunitarias[32].

Evolución y deriva natural

Un tercer paralelo crítico al que he llegado naturalmente en estos años es que para poder dar sentido a un sistema autónomo, el pensamiento evolutivo debe dejar atrás una visión neodarwinista que entiende la selección como fuente instructiva de modificaciones históricas. La nueva visión que emerge en consonancia con todo un movimiento renovador en biología evolutiva se hace cargo de las capacidades auto organizativas intrínsecas del organismo a nivel genotípico (el genoma es una red compleja) y durante el desarrollo embriológico (el desarrollo es una transformación integrada, y no la expresión de caracteres). Además, el acoplamiento estructural con el medio se realiza, no sólo a nivel del individuo, sino también a varios otros niveles, tanto celular como poblacional, y sobre la base de ciclos completos de vida. El énfasis en la constitución interna y los múltiples niveles de imbricación del ciclo de vida de todo organismo llevan a cambiar la visión de la selección natural clásica a una que puede designarse como *deriva natural*. En la evolución

32. Huetz, F.; Jacquemart, F.; Peña-Rossi, C.; Varela, F.; Coutinho, A., "Autoimmunity: the rnoving boundaries between physiology and pathology", en J. *Autoimmunity*, 1988, 1: 507-518; Stewart, J., Varela, F.; Coutinho, A., "The relationship between connectivity and tolerance as revealed by computer simulation of the immune network: some lessons for an understanding of autoimmunity", en *J. Autoimmunity.*, 1989, 2 (Supplement): 15-23; Lundqvist, 1.; Coutinho, A.; Varela, F.; Holmberg, D., "Evidence for the functional dynamics in an antibody network", en *Proc. Nat. Acad. Sci.*, EE.UU., 1989, 86: 5074-5078; Varela, F., Andersen, A.; Sundblad, A.; Holmberg, D.; Kazawhkine, M.; Coutinho, A., "The population dynamics of natural antibodies in normal and autoimmune individuals", en *Proc. Nat. Acad. Sci.*, EE.UU., 1991, 88: 5917-5921; Dietrich, G., Varela, F.; Kzcí, M., "Manipulating the human immune network with ig", en *Europa J. Immunol,* 1993, 23: 2945-2950.

como deriva natural, la selección aparece como condiciones de borde que deben ser satisfechas, pero al interior de las cuales el camino genotípico y fenotípico de un organismo se funda en su cierre operacional. Las consecuencias de todo esto son, por cierto muy importantes, pero no es éste el lugar para entrar en detalles[33].

Enacción y cognición

Una de las críticas que debe hacerse a esta obra (así como a mi libro del 79), es que la crítica de la representación como guía del fenómeno cognitivo es reemplazada por una alternativa débil: lo externo como mera perturbación de la actividad generada por el cierre operacional, que el organismo interpreta ya sea a nivel celular, inmunitario o neuronal. Reemplazar la noción de *input-output* por la de acoplamiento estructural fue un paso importante en la buena dirección porque evita la trampa del lenguaje clásico de hacer del organismo un sistema de procesamiento de información. Pero es una formulación débil porque no propone una alternativa constructiva al dejar la interacción en la bruma de una mera perturbación. A menudo se ha hecho la crítica que la *autopoiesis,* tal y como está expuesta en este libro, lleva a una posición solipsista. Por lo que acabo de decir, yo pienso que ésta es una crítica que tiene un cierto mérito. La tentación de una lectura solipsista de estas ideas deriva de que la noción de perturbación en el acoplamiento estructural no toma adecuadamente en cuenta las regularidades emergentes de una h*istoria* de interacción en donde el dominio cognitivo no se constituye ni internamente (de un modo que lleva efectivamente al solipsismo), ni externamente (como lo quiere el pensamiento representacionista tradicional). En estos últimos años he desarrollado una alternativa explícita que evita estos dos escollos, haciendo de la reciprocidad histórica la clave de una *codefinición* entre un sistema 'autónomo' y su entorno. Es lo que propongo

33. El término deriva natural aparece por primera vez en *El árbol del conocimiento, op. cit,* capítulo 7. Para un desarrollo más extenso véase Varela, F.; Thompson, E.; Roscii, E., 1992, *The embodied mind,* MIT Press, Cambridge, 1991, chap. 9.

llamar el punto de vista de la *enacción* en la biología y ciencias cognitivas[34]. Enacción es un neologismo, inspirado del inglés corriente en vez del griego como lo es la *autopoiesis.* Corrientemente enacción se usa en el sentido de traer a la mano o hacer emerger, que es lo que me interesa destacar. La prueba de fuego de este punto de vista es que ha permitido una reconstitución detallada y meticulosa de un fenómeno que puede verse como caso ejemplar: la visión de colores[35].

Conciencia y fenomenología

La última pieza del puzzle a considerar aquí es la relación de estas ideas sobre el ser vivo en sus varias dimensiones y el fenómeno de lo mental y la conciencia, entendida como experiencia vivida. De hecho el problema de la conciencia vuelve a estar a la orden del día hoy: no se puede hacer neurociencias sin dar una respuesta, aunque sea implícita, a esta cuestión. Desde mi punto de vista, la crítica del conocimiento que acompaña a la a*utopoiesis y* que se continúa al poner en evidencia el rol central de la interpretación y de la autonomía de los seres vivos, es clave. Sin embargo ese nuevo análisis biológico requiere un complemento de una *disciplina fenomenológica* de la experiencia. Esta exigencia de una disciplina adecuada es un punto ciego en nuestro discurso científico que se contenta con suponer que la vida mental aparece como algo evidente y accesible a los humanos, lo que es un error radical. Esta crítica se aplica a la cuasi totalidad de los escritos actuales sobre la conciencia y aún a lo que decimos en el *Árbol del conocimiento,* donde se reclama un rol para la experiencia (es el comienzo y el fin del camino del itinerario conceptual que el libro recorre), pero no se ve más allá en el esfuerzo por buscar una metodología explícita para poder

34. Varela, F., *Conocer: las ciencias cognitivas,* Gedisa, Barcelona, 1990. Varela, F., "Organism: A meshwork of selfless selves", en Tauber, F. (ed.), *Organism and the Origin of Self*. Kluwer Assoc., Dordrecht, 1991, pp. 79-107.

35. Thompson, E.; Palacios, A.; Varela, F., "Ways of coloring: comparative color vision as a case study in cognitive science", en *Beh. Brain Scien,* 1991, 15: 1-75.

tematizar esa experiencia. En este sentido he vuelto a recobrar mi antiguo interés en la fenomenología europea, y en particular en Merleau-Ponty, como el antecedente más cercano a mi posición enactiva, aunque añado a esa tradición europea, un sostenido interés en una disciplina de la experiencia que encarna la tradición de meditación budista a lo largo de 20 siglos. Esta *circulación* necesaria entre experiencia y ciencias cognitivas es el tema de mi libro más reciente, quizá el escrito que más me ha exigido en un esfuerzo de síntesis, para lograr poner lado a lado la enacción y la experiencia entendida como una disciplina rigurosa[36]. Mi interés actual está centrado sobre todo en esta reciprocidad: externalidad de la operación cognitiva y fenomenología de la experiencia vivida. En el laboratorio, las nuevas técnicas de imágenes cerebrales permiten explorar todo esto de manera empírica y precisa. Es el comienzo de una nueva ciencia bio-fenomelógica donde falta casi todo por hacer.

CODA

Como se ve, desde mi perspectiva de 1994, la *autopoiesis* no encarna sola una nueva visión de la vida y el conocimiento. A su lado aparecen otras nociones igualmente importantes, tales como el cierre operacional, la enacción, la deriva natural y la metodología fenomenológica. Los referentes empíricos se extienden consecuentemente en nuevos programas de investigación detallados, sean las redes de linfocitos, la marcha de los insectos, o las imágenes cerebrales. Se trata de un edificio de nuevas concepciones epistemológicas y de resultados empíricos que tiene gran amplitud y se sostiene con rigor. Han sido 20 años productivos en los que la etapa de formulación de la *autopoiesis* marca, en retrospectiva, un hito importante como debería ser evidente para el lector que ha tenido la paciencia de seguirme hasta aquí. Este Prefacio es también para mí la ocasión de agradecerle públicamente una vez más a Humberto

---

36. Varela, F.; Thompson, E.; *Roscii, E.; The embodied mind: cognitive science and human experience,* MIT Press, Cambridge, 1992. Hay traducción española en Editorial Gedisa, Barcelona.

Maturana el haber sido mi primer maestro en el juego de la ciencia, y luego colega en la aventura de lo nuevo. Pocos son los que tienen su coraje y lucidez como pensador; si nuestros encuentros no hubieran sido, mi vida tendría una dimensión de menos.

Pero si esta construcción lenta y sostenida, llena de *corsi y ricorsi* como toda creación intelectual, tiene hoy día viabilidad científica, es porque se inserta en una sensibilidad histórica que la *autopoiesis* intuía ya en 1973. Como decía al comienzo de este Prefacio, no hay creaciones personales desprovistas de una inserción: el que una idea tenga un impacto es un hecho histórico y no una aventura personal o un asunto de "tener razón". La *autopoiesis* sigue siendo un buen ejemplo de alineamiento con algo que sólo aparece más claramente configurado en varios dominios del quehacer cultural humano y que indicaba yo con el nombre de un giro ontológico. Es decir una progresiva mutación del pensamiento que termina con la larga dominancia del espacio social del cartesianismo y que se abre a la conciencia aguda de que el hombre y la vida son las condiciones de posibilidad de la significación y de los mundos en los que vivimos. Que conocer, hacer y vivir no son cosas separables y que la realidad y nuestra identidad transitoria son *partners* de una danza constructiva. Esa tendencia que designo como un giro ontológico no es una moda de filósofos, sino que se refleja en la vida de todos. Entramos en una nueva época de fluidez y flexibilidad que trae detrás la necesidad de una reflexión acerca de la manera de como los hombres hacen los mundos donde viven, y no los encuentran ya hechos como una referencia permanente.

Las consecuencias éticas de este darse cuenta son importantes, y en todo caso suficientemente actuales para que merezcan una discusión más extensa que la que puedo hacer aquí. Pero insisto sobre este punto porque la ocasión de escribir este prefacio que celebra estos 20 años, sería tristemente dilapidada si no llegara a comunicar al lector la importancia de expandir el horizonte para considerar el carácter profundamente social y estético en donde esta idea se inserta, más allá de la ciencia y la biología, y más allá de las personas que figuran como autores. En ese sentido este pequeño

libro no ha perdido vigencia y todavía puede leerse con provecho. En definitiva una invención científica en cualquier campo requiere actores que sean sensibles a las anomalías que siempre nos rodean. Esas anomalías deben ser mantenidas en un estado de suspensión y cultivo mientras se es capaz de encontrar una expresión alternativa que reformule la anomalía como un problema central, tal y como la *autopoiesis* pone la autonomía al centro del problema de la vida y el conocer. Quizá el caso de la *autopoiesis*, en el que me ha tocado la fortuna de ser partícipe, pueda servir para ilustrar esta dinámica de la innovación y contribuir así a que el futuro de la ciencia en Chile sea responsable de cultivar sus sensibilidades propias y no sea un eco de tendencias de otras latitudes.

"NINGÚN LUGAR ESTA AQUÍ O ESTA AHÍ"

Este diálogo con Cristián Warnken tuvo, me parece, un nivel de inspiración y de altura que le hace honor al nombre de la emisión donde se difundió: "La belleza de pensar". *Los temas tocados son varios y diversos, y me parecieron la mejor manera de cerrar esta serie, con su modo de pasar de un tema al otro, dejando múltiples ventanas abiertas a lo que será siempre el deleite de pensar la vida como* todo *el fenómeno.*

Cristián Warnken: El poeta chileno Oscar Hahn, en su poema "ningún lugar está aquí o está ahí", afirma:

"Ningún lugar está aquí o está ahí
Todo lugar es proyectado desde adentro
Todo lugar es superpuesto en el espacio
Ahora estoy echando un lugar para fuera
Estoy tratando de ponerlo encima de ahí
Encima del espacio donde no estás
A ver si de tanto hacer fuerzas y de tanto hacer fuerzas
Te apareces ahí sonriente otra vez
Aparécete ahí aparécete sin miedo
Y desde afuera avanza hacia aquí
Y has harta fuerza, harta fuerza
A ver si yo me aparezco otra vez si aparezco otra vez
Si reaparecemos los dos tomados de la mano
En el espacio donde coinciden todos nuestros lugares.

(Varela, F. (1997), La Belleza de Pensar (entrevista de Cristián Warnken), ARTV, Santiago de Chile).

Siempre me ha gustado este poema de Oscar Hahn, y sobre todo estos versos: "ningún lugar está aquí o está ahí/todo lugar es proyectado desde adentro". Lo primero que quisiera preguntarte es si estás de acuerdo con lo que afirma el poeta Oscar Hahn.

Francisco Varela: Creo que Oscar Hahn dice que lo que está afuera no está afuera ni está adentro, y eso es atractivo. Si tú me pones frente a la perspectiva clásica de decir: "el espacio es una cosa física, independiente del ser vivo" y si me hago una imagen interna, eso es demostrablemente falso.

CW: ¿Idealismo?

FV: No, es un realismo, un empiricismo de relativo primitivismo filosófico y además demostrablemente falso; voy a decir luego por qué. Por otro lado, si tú me pones frente a la alternativa de decir que yo desde adentro proyecto cualquier cosa...

CW: ... eso es el idealismo...

FV: ... eso es más bien el idealismo o lo que hoy día se llama constructivismo, al que tampoco adscribo para nada. Lo bonito es encontrar una posición intermedia, la vía del medio, del conocimiento, en la cual, por ejemplo, el espacio no es ni afuera ni adentro, es decir, que el espacio que está frente a mí, yo no lo puedo separar de los gestos que hago para apuntarlo y cuando digo apuntarlo, literalmente quiero decir apuntarlo (*apunta con el dedo*). Algunos, por cierto, dicen: "¡Ah! Mira qué interesante, eso me recuerda la filosofía, por ejemplo, de Merleau-Ponty o de Husserl", lo que es cierto, pues es un tema favorito de los fenomenólogos. Sin embargo, vemos hoy en día en la neurociencia que esas mismas ideas uno las puede ver encarnadas en el cuerpo, por ejemplo, cuando digo: "Esa lámpara que está allí no es separable de mi brazo" y eso es porque, hoy por hoy, uno puede demostrar que la percepción de mi brazo, la propiocepción de mi sentimiento de la posición del brazo, no es separable de que yo pueda señalar un allá, y ahí es donde Oscar Hahn tiene razón, o sea que el espacio allá lo estoy poniendo, pero no lo estoy poniendo de una manera arbitraria, lo estoy poniendo justamente en este gesto, en esta gestualidad que es

la base de la vida animal; la vida del conocimiento es la raíz; en el fondo es esta capacidad de moverse, de tener esta generación interna de una vitalidad que termina en un gesto y en ese gesto se constituye la percepción y allí, tanto ésta como la acción, se hacen una sola cosa y, por lo tanto, el espacio no está ni adentro ni afuera, está en este movimiento, en este ir agarrándose de lo que no está y lo voy poniendo y después lo voy sacando. ¿Es muy bonito, no?

CW: Es muy plástico.

FV : Exactamente.

CW: Esto tiene una serie de implicancias, ¿no? Hasta en la vida personal, si uno toma conciencia de lo que tú estas diciendo, tengo la sensación de que la manera de aproximarse al mundo comienza a ser distinta, no sólo filosóficamente, sino también en términos reales. ¿Estás de acuerdo con eso?

FV: Absolutamente. Ya que estamos en eso, entremos en el punto. Cuando yo vivo de esa manera de reentenderme —porque en el fondo es una manera distinta de estar parado en el mundo—, lo que es válido para el espacio lo puedo decir también para el tiempo, para las emociones, para la memoria. Uno dice, "¡Vaya! esto cambia todo", yo no puedo, por ejemplo, comportarme como si las cosas vinieran de fuera del mundo tan simplemente, o que el otro estuviera allá afuera y yo acá adentro. Sin embargo, una de las cosas interesantes que tiene la mente, para usar una palabra clásica, es que el saber no cambia el ser. Yo puedo saber esto, intelectualmente comprenderlo e incluso poder demostrarlo, mostrarte los registros o mostrarte los argumentos filosóficos. Sin embargo, eso no cambia mi modo de ser cotidiano, en el cual yo me sigo comportando de esa manera ingenua. En esa filosofía espontánea, dirían los filósofos, en la cual el mundo está allá afuera y me llega, o el otro está allá afuera y me ataca, o yo estoy aquí y soy distinto de lo otro. Esa actitud natural está anclada y sobrepasa el saberlo o no saberlo; allí hay una gran trampa, entre saber que eso es así y comportarse de otra manera. Medir esto pasa, no sólo por la comprensión, sino por un trabajo, una disciplina interna de cambio, que tiene que

ver con la observación muy minuciosa, la fenomenología de la experiencia.

CW: Parece que es un rasgo del pensamiento occidental separar la realidad en dos dimensiones o en dos realidades: sujeto-objeto, por ejemplo; o afirmar que casi todo viene desde afuera, es quizá uno de los rasgos que nos ha caracterizado y nos caracteriza como como manera de ser profunda, occidental.

FV: Si tú me permites, creo que eso es un mito. Ese 'cartesianismo occidental'—del que Descartes no tiene la culpa— o tendencia espontánea al dualismo, que es de eso de lo que estamos hablando, no es ni occidental ni oriental, ni antiguo ni moderno; es constitutivo en los animales, en los niños; es constitutivo a todas las culturas. En todas partes ha habido gente capaz de desmontar esa idea. Por ejemplo, en occidente podemos citar la fenomenología de Edmond Husserl y de Merleau-Ponty —y yo soy un gran admirador de ellos, pero ciertamente de este último menos—. Husserl es una de las cumbres del pensamiento humano; él vio con gran claridad esto de lo que estamos hablando, sin embargo, en la ciencia podemos encontrar este mismo argumento. Sin embargo, entender y encontrar maneras para realmente encarnar eso, deshaciendo, desmontando esa actitud natural, es otra cosa, y ahí es donde yo siento que occidente ha perdido trazas de esas prácticas de transformación, para utilizar el término de Michel Foucault; prácticas de transformación humana, prácticas de trabajo con la experiencia humana, pero además, no tiene conciencia de que las perdió, salvo excepciones. En cambio una de las cosas hermosas que nos da la tradición budista no es su mayor o menor ingenuidad, ya que los budistas son tan dualistas como nosotros; tienen el mismo problema, sino su tradición filosófica y su fenomenología de la experiencia. El budismo nunca ha perdido de vista esto; ha sido una obsesión la de cultivar esas prácticas; el budismo es una verdadera mina de oro de observaciones acumuladas, de pragmática, de *know-hows*, que es lo que a mí me entusiasma de esa tradición y no ese mito occidental de que los orientales son todo '*om*'.

CW: ¿Y en occidente, aparte de la ciencia y la filosofía, no hay también una tradición homologable a esa, quizá en el arte o en la literatura o la poesía? Pienso en los románticos, en Novalis, en muchos poetas, poetas místicos, una tradición de pensamiento paralela, de alguna manera, a la filosofía.

FV: Sin la menor duda. Por ejemplo, Kafka es un gran literato, por lo menos como yo lo leo. Entendió eso hasta el fondo, pero lo que la literatura occidental ha logrado ahí, en ese sentido, es como la ciencia o como la filosofía; ahí hay un saber y un entender, pero que no contiene, salvo excepciones, que no es el caso de Kafka, la práctica que va con eso. Hay otros literatos o filósofos que han intentado montar una práctica detrás, pero son muy raros, estoy pensando, por ejemplo, en lo que fue el movimiento de los surrealistas en Francia, ellos intentaron montar una práctica pero son intentos que se quedan ahí y no llegan a encarnarse en una continuidad; en cambio, en la tradición budista, por ejemplo, hay veinticinco siglos, veinticinco siglos de experiencia. Imagínate lo que sería, por ejemplo, una práctica literaria del estilo de los surrealistas o de los románticos, sistemática, cultivada durante veinticinco siglos, ¡eso podría ser maravilloso!

CW: ¿Qué nos hemos perdido los occidentales al no tener este eslabón, este paso del saber a la práctica? ¿ Qué males derivan de eso?

FV: Creo que los dos males más populares del momento: el nihilismo o bien el dogmatismo, dos caras de una misma moneda, o sea, o bien es necesario afincarse en una especie de sistema de creencias muy sólido, lo que es típicamente religioso hoy por hoy, pero que puede ser también cientista no científico...

CW: ... un fundamentalismo científico...

FV: Un fundamentalismo científico, que es más raro o puede ser también un fundamentalismo del tipo de eficacia productiva, un neoliberalismo militante a la Milton Friedmann, por ejemplo, Y si no, de manera más larvada, ese nihilismo que hace que en Chile haya más suicidios que en ninguna parte del mundo y que se

consuman prácticamente tantos tranquilizantes como en el resto de Latinoamérica. De todo esto último me enteré cuando fui invitado a dar una conferencia en el congreso mundial de psiquiatría y neurología. No es el nihilismo reflexivo nietzschiano, es un nihilismo vivencial y esos son los males en que uno es la contracara del otro.

CW: Ahora; tú hablas ahí de una práctica. Occidente hizo el otro camino: ¿cuál sería?

FV : Digamos que perdió una opción. La tuvo en la Edad Media, pero la perdió.

CW: En la ciencia contemporánea de hoy, en el mundo científico, que es donde se están haciendo grandes preguntas y donde se está haciendo filosofía de alguna manera, quiérase o no, este planteamiento tuyo, ¿es minoritario, de un grupo de "locos", que están en algunas clases, en algunos lugares o es un planteamiento que está ya socavando o penetrando o permeando la ciencia contemporánea? Porque la ciencia que uno conoce es la ciencia decimonónica, una ciencia más soberbia.

FV: Bueno, no hay que preguntarle nunca a los locos si lo están porque te van a decir que no... Una de las cosas maravillosas que ha hecho la ciencia en occidente es que le ha devuelto a occidente justamente una práctica. La ciencia en occidente es donde está la práctica, es lo que le da su fundamento creativo, cultivado y sostenido ya por cuatro siglos, digamos tres siglos, de una manera u otra, y ciertamente en los últimos ciento cincuenta años. Esa es la gran flor de occidente a mis ojos, por eso es que yo soy un científico orgullosos de serlo y que reivindico mi pertenencia al clan. ¿Por qué?, porque ese método, como todos los fundamentos metodológicos sostenidos y mantenidos, que no están fundados sobre un criterio de autoridad o un criterio de revelación o de palabra revelada, contiene esa maravillosa capacidad de irse transformando, con tiras y aflojas. Nadie pretende que la ciencia es pura apertura, sólo autocrítica; es falso, pero contiene ese elemento de creatividad autocrítica que la lleva justamente a sobrepasar esas limitaciones de tipo decimonónicas, y de eso todos

estamos muy conscientes. Si tomas este libro que tienes aquí sobre los nuevos paradigmas o si vas a una conferencia científica, existen hoy en día en las ciencias que se ocupan de lo mental —las neurociencias cognitivas— dos actitudes, una de las cuales tiende todavía a tirar hacia una posición, para decirlo rápido, muy reduccionista. Es decir, la conciencia de la experiencia es un hiperfenómeno, denme tiempo y voy a encontrar la molécula, la célula, el circuito, te das cuenta de eso y se acabó el problema...

CW: ¿Podría decirse que es casi un reduccionismo físico?

FV: Un reduccionismo físico-biológico, el reduccionismo físico a estas alturas es demasiado burdo, así es que ya nadie lo mantiene, pero un reduccionismo como, para usar el nombre de un libro de un neurocientífico francés, Jean-Pierre Changeux, *El Hombre Neuronal*, no es el hombre-máquina, es el hombre neuronal... Pero lo interesante que tiene, y por eso hago referencia a esa continuidad del método y eso de contener una especie de crítica interna, de creatividad interna, es que cuando tú tienes una disciplina como las neurociencias o las ciencias cognitivas, que estudian la mente, ese objeto, de alguna manera no se te puede arrancar completamente, entonces tú no puedes hacerte el desentendido, pretender que la experiencia y la conciencia no existen, uno lo puede negar por un tiempo, pero rápidamente eso hace agua. Si yo estudio los cuarks o los leptones o las galaxias o el estado sólido, puedo desentenderme de eso completamente; no estoy comprometido a encontrarme al final de mi investigación con este tremendo signo de interrogación: "¿Y usted dónde está, señor?" Eso no aparece en la física.

CW: ¿Ni siquiera en la física cuántica?

FV: Sólo de manera muy abstracta, que es donde está el observador; es una manera muy elegante; en cambio aquí nos encontramos con este "¿dónde está usted, señor?", el que sufre, el que duerme, ese es el que me interesa, ese es jugoso, el otro observador es una especie de ojito pegado en la pared, muy conveniente. Los mecánicos cuánticos y los físicos creen que están hablando de la conciencia del observador, lo que no es falso, pero sí cómodo, pues nos libera de

todo el lado enjundioso, donde está la verdad vivencial, donde yo me encuentro, porque nadie se identifica con un observador que mira moleculitas en una pantalla en un experimento de mecánica cuántica, no somos eso, sin embargo sí somos los que pensamos, los que sufrimos, los que sentimos, los que nos dormimos, los que soñamos...

CW: Hay grandes preguntas sobre este tema de la mente que tal vez nunca van a ser resueltas. ¿O crees tú que que no es así o bien que haya que aceptar que va a haber siempre preguntas sin contestar?

FV: A eso me quería referir cuando te decía que hay toda una corriente reduccionista que sigue siendo así; sin embargo, como en esta disciplina tú no te puedes escapar a encontrarte contigo al fondo del espejo, eso ha generado de manera recurrente una especie de re-emergencia del problema de la conciencia, que tiene fases. En cambio, cuando yo era estudiante, hablar de la conciencia en círculos científicos era mal visto, mala educación. Volviendo a lo que te decía hace poco, esto viene por fases y en los últimos cuatro o cinco años viene muy fuerte, aunque desde hace diez años se viene instalando, de a poquitito. Es un tema que se ha vuelto a constituir como un tema legítimo. El año pasado, por ejemplo, hubo un gran congreso en Estados Unidos llamado : «Hacia una ciencia de la conciencia», y ahí se debatían temas como juntar la neurona con la experiencia, así brutalmente, evidentemente hay opciones; hay aquellos que dicen: estos dos jamás se encontrarán y uno tiene que quedarse con esta especie de *flip flop*, hay otros como yo, y no soy el único, aunque somos minoritarios, que sostienen que estas dos cosas tienen una relación de circulación mutua, de puente. Si yo tomo este vaso (t*oma un vaso y bebe*) tengo la experiencia del agua, pero al mismo tiempo nadie ni nada me impide mirar lo que pasa con el circuito que me hace mover el brazo, ahí hay una relación entre dos cosas. ¿Por qué voy a sentir yo que son contradictorias? Evidentemente el tema, allí, no es que sean o no sean contradictorias, sino más bien qué hay que hacer para encontrar los puntos de contacto en que esas implicaciones se dan y hoy por

hoy eso lo podemos hacer, lo podemos hacer gracias a dos cosas y ese es mi trabajo actual. Primero, hay que tener los métodos que nos permitan mirar este cuerpo, este sistema nervioso en acción intocado, intacto, métodos que se llaman de imágenes cerebrales. En el laboratorio, en París, tenemos una máquina maravillosa que se llama de resonancia magnética, donde tú puedes poner a un sujeto, intocado y le puedes decir por ejemplo: beba agua...

CW: Y ver qué pasa cuando bebe agua...

FV:... y al mismo tiempo, recoger una imagen de la dinámica interna del sistema nervioso a distancia, gracias al hecho de que hay estas señales paramagnéticas de los movimientos moleculares que, con mucho trabajo y con mucho cálculo —ahí es donde está el 99 por ciento del sudor de la investigación, pero que importa— tú obtienes esas imágenes en acción. Al mismo tiempo, puedes tener a un sujeto que toma el agua, y ahí es donde yo vuelvo a mi explicación metodológica, ese sujeto puede ser un sujeto vulgaris que dice "estoy tomando agua" o bien puede ser un experto que es capaz de discriminar momentos muy finos de la experiencia, por ejemplo: en este momento estoy trabajando...

CW: Se podría hacer con el vino también (*risas*), alguien que supiera distinguir las diferentes cosas que ocurren dentro nuestra boca cuando bebemos vino.

FV: Excelente ejemplo: Imagínate hacer ese experimento del agua, con vino entre dos sujetos; uno, como yo, que nunca toma vino y un enólogo, que lleva treinta o cuarenta años en esa actividad y que es capaz de discriminar respecto de las cepas o del año de un determinado vino. Este experimento es equivalente a lo ocurrido en neurociencia donde se ha trabajado como si la experiencia fuera una sola cosa y simple.

CW : ¿Eso es conductismo?

FV: No, yo diría que es una especie de actitud natural. Decir que tener experiencia implica ser experto es falso. Tener un cuerpo no implica ser médico; si yo escucho sonidos, no soy músico; si tomo

vino, no soy enólogo. La maestría se adquiere con un entrenamiento largo. Y eso es válido respecto del tiempo o de las emociones. Esa es la fenomenología de la experiencia, eso es lo que aprendemos, por ejemplo, de la tradición budista; eso es lo que me interesa de la fenomenología y eso es lo que podemos hacer en el laboratorio y que permite establecer esos puentes porque tiene un material empírico, datos fenomenológicos suficientemente finos; de lo contrario, el desbalance es muy grande entre lo refinado de los métodos externos y lo, un tanto brutal, por decir lo menos de los métodos internos.

CW: ¿Cuál de todos los misterios de la conciencia te parece a ti él más apasionante o el que más te impulsa a seguir trabajando en determinada dirección?

FV: El tema que para mí sigue teniendo el nervio de todo esto es algo que al mismo tiempo es muy simple y muy profundo, me refiero simplemente a la capacidad de darse cuenta. Es decir, de pronto estás haciendo algo y te haces esa curiosa reflexión, ese curioso gesto que consiste que, en vez de que la vida perceptiva mental esté dirigida hacia fuera, das vuelta la dirección de la atención y estás en ese estado de poder acoger de manera espaciosa lo que es esta experiencia, sin pegarte a ella y, sin embargo, seguir diciendo y haciendo las mismas cosas.

CW: Es el desapego, el "desasimiento", le llamaba santa Teresa...

FV: Exactamente o para usar un término de Heidegger, "*Esgelassenhein*". Husserl habla de "*epoge*" o de suspensión, los budistas hablan justamente de "suspensión" o de reducción, lo puedes poner entre paréntesis; todas esas son fórmulas que apuntan exactamente a lo mismo. Ese gesto, que todos lo hacen y ocurre espontáneamente, la vida te impulsa a eso: tú vas por la calle, de repente alguien te toca un bocinazo, pero lo curioso es que uno pasa de largo, y seguimos pensando en la reunión, etc. Entonces, para mi la práctica fenomenológica es cultivar esa capacidad, pero de manera radical, es decir, cultivarlo de tal modo que tú haces la reducción hasta el punto en que tu conciencia no tiene contenido.

CW: ¿Y eso para qué?, ¿ para conseguir paz, felicidad, plenitud.?

FV: Para conseguir nada, simplemente porque está ahí. A mí ni la parte de la paz ni de la salvación me interesan pues soy un epistemólogo; a mí me interesa conocer, quiero ver lo que hay ahí. Ahora, es cierto que cuando tú mantienes ese estado de suspensión, de conciencia sin objeto, es cierto que tiene una espaciosidad que te toca de una manera existencial muy profunda; no es la paz, pero te llega al fondo, te sientes en casa... Es otra cosa, una especie de nirvana...

CW: Si entiendo. bien esa espaciosidad que tú estás definiendo, que tiene varios nombres, ¿sería eso lo que nos definiría como propiamente humanos?

FV: Sí, y lo que me gusta de esta manera de enfocarlo es que, como estamos hablando de una serie de etapas y de caminos para cultivar esa capacidad, quiere decir que al mismo tiempo yo lo puedo estudiar, puedo hacer mi "neuro-fenomenología", que es así como yo llamo a esta disciplina. Yo soy un "neurofenomenólogo".

CW: Suena pomposo... (*risas*) Pero suena bonito también... (*risas*)

FV: Pensé que sonaba bonito, no más, pero en fin...(risas)... Hablo de neuro-fenomenólogía porque mantiene dos cosas juntas, ese gesto de suspensión, de *epogé*, de desasimiento, como tú decías, citando a santa Teresa, de *gelassenheit*. Cómo ocurre eso, cómo se da, cuáles son sus bases, sus límites, sus fundamentos, bajo qué condiciones se dificulta, qué tipo de problemas de orden constitucional, bioquímicos, de circuitería. Entramos incluso al nivel de los cables y los jugos, como digo yo...¿por qué no? Y al mismo tiempo mirarlo desde el punto de vista de aquel que lo experimenta, que lo puede experimentar a distintos niveles y ahí es donde tenemos otra vez el abanico entre el enólogo experto y el *naïf*. Eso es lo que estoy haciendo, estoy trabajando con todo ese espectro de competencias sobre sí mismo. Y creo que, aunque parece pretencioso decir que eso es lo que caracteriza lo humano, ciertamente hay algo ahí muy cercano a lo que digo.

CW: Me gustó eso de 'sentirse en casa'... es verdad que algunas veces, en ciertos momentos del día es como una intimidad, como un refugio, un espacio en donde uno se siente en plenitud

FV: Es una sensación de relajación profunda, es una cosa incondicional que al mismo tiempo está siempre allí, yo no tengo que ganármela, no tengo que ir a ninguna parte. Es ciertamente un estado mental, pero al mismo tiempo toca la vida espiritual del hombre. Entonces no puedo pensar, por mucho que mis colegas me miren de repente con cara de desaprobación, que uno pueda estudiar la vida, la mente y el cerebro, sin eventualmente incluir lo que esa dimensión de la experiencia nos rinde a la tradición espiritual del hombre.

CW: ¿Qué es lo que sería esa dimensión espiritual? Me gustó escucharlo en boca de un de un científico...

FV: Para mí la vida espiritual es mantener apasionadamente presente esta pregunta: Por qué no estoy en casa y qué tengo que hacer, no a quién tengo que creer o qué repeticiones tengo que hacer, sino cómo me doy los instrumentos para transformar mi ignorancia, para poder entender qué es lo que me separa de algo. Y, por lo tanto, para mi vida espiritual, incluido este camino de cultivar esa capacidad de ser, es muy importante lo que yo he aprendido con esta gente.

CW: Tú has intentado en ese bellísimo libro, *Un puente para dos Miradas**, de tender un puente entre dos tradiciones —el budismo y la ciencia occidental— en relación con el tema de la cognición, de la percepción, de la mente. ¿Por qué el budismo, por qué no, por ejemplo, dentro de occidente, la tradición mística cristiana, por qué no el judaísmo? El budismo es hoy, para muchos, una moda, se habla mucho del budismo, todo el mundo en Hollywood se ha hecho budista...¿Por qué el budismo y no otra tradición?

FV: Parto por señalarte que eso lo comenzamos nosotros mucho antes de este efecto... Yo conocí al Dalai Lama en el 83, cuando, sin ser un desconocido, no era un personaje *vedette*; y mi interés

---

* *Un puente para dos miradas. Conversaciones con el Dalai Lama sobre las ciencias de la mente.* F. Varela y J. Hayward, editores. Dolmen Ediciones, Santiago de Chile, 1987.

por la tradición budista viene del año 74, cuando no era para nada una moda. Y si esto ocurre, no me parece ni bien ni mal, pero no es mi problema. No me siento comprometido con ese efecto.

¿Por qué el budismo? Porque es la única tradición cultural sobre este planeta —estoy hablando de occidente y oriente—, que ha mantenido una tradición sostenida de análisis fenomenológico de la experiencia bajo una continuidad de método con gran perseverancia y con un gran rigor para acumular un corpus filosófico, de observación, metodológico, extraordinario. Entonces lo menos que podemos hacer es escuchar eso. No se trata tampoco de adoptarlo de lleno, pero sería ridículo rechazarlo sin fundamento, es como si viviéramos en algún apartado lugar de otro continente y de pronto nos diéramos cuenta de que las matemáticas son importantes y que ellas se hicieron hace veinte siglos en Europa y Estados Unidos y las rechazáramos por el solo hecho de ser occidental. Sería un tanto estúpido ¿no? Tú mencionabas la mística occidental y allí hay bellísimos ejemplos, Renán o Eckhart, por ejemplo. Tú lees a éste y dices: "Este hombre *sabe*". Sabemos que tenían prácticas, por los estudios que se han hecho; sin embargo, de eso hay pocas huellas.

CW: No hay tradición, quizá.

FV: No hay cultivo de esas tradiciones, no hay escuela, no hay líneas continuadas. Entonces es como si alguien, de pronto, hubiera descubierto ciertos capítulos de la ciencia, pero los pasos intermedios se hubieran perdido, de modo que hay que reinventar la receta. Tenemos la evidencia de que fue, pero eso es lo que occidente ha perdido.

CW: A lo mejor si el cristianismo hubiera tomado esa huella, otra habría sido la historia del cristianismo en occidente.

FV: No me cabe la menor duda. Tú lees autores de los siglos X, XI, XII, que no tienen nada que envidiarle a Heidegger o a Husserl, o a los tratados de psicología experimental, son obras monstruosamente majestuosas.

CW: Tú conociste al Dalai Lama personalmente y él hace preguntas muy incisivas que de repente a uno lo sorprenden, como si ya hubiera

estudiado este tema, con otro lenguaje. ¿Se podría decir, entonces, que el budismo, de alguna manera, es también una ciencia que se desarrolló paralela a la de occidente, con otro lenguaje, con otra formulación?

FV: La manera correcta de describirlo es —pienso— como una fenomenología aplicada, que tiene el rigor en la filosofía y el enraizamiento en la práctica y, desde ese punto de vista, tiene una interlocución muy fácil con la ciencia porque es la misma cosa, tienen en común el terreno de la experiencia, el terreno del examen, el terreno de la autocrítica. Más que una ciencia, porque es un pocoo pretencioso decir que lo es, si por ciencia entendemos lo específico que tiene la ciencia de occidente. Ellos no pretenden una ciencia, pero sí una disciplina, un rigor.

CW: La experiencia que tú viviste aquí es una experiencia, privilegiada, un lujo ¿no? Haberse "encerrado" ahí un grupo de científicos occidentales de alto nivel y el Dalai Lama y otras personas. ¿Qué fue para ti esa experiencia?

FV: Lo primero que hay que entender que para mí esa experiencia no es aislada. Yo comencé a interesarme en el budismo en el año 74 y desde entonces he conocido mucha gente, hombres y mujeres extraordinarios de varias procedencias dentro de la tradición budista, ya sea, de Sri Lanka o de origen más bien de tendencia china, japonesa o tibetana. El Dalai Lama es extraordinario, pero no es el único. En segundo lugar, esa fue una experiencia que tuvimos, pero hubo otros encuentros antes y siguieron otros después. Es todo un panorama que no puedo separar de ese contexto. Yo me he dado cuenta lentamente de lo que hay en el budismo; es una mina de oro y uno tiene acceso a ella a través de lecturas Lo maravilloso de esas experiencias es que te dan un acceso en vivo y en directo, así como le ocurre al experto, al enólogo (*risas*), porque él lo es, igual que otros que te hablan de la tradición con una voz que está encarnada, que está viva, que no es cosa de repetir una receta, o sea, no tienes el menú, tú ves cocer el pan y después te lo dan. Esa es un poco la cosa, son experiencias muy especiales y yo me siento muy privilegiado de haberlas tenido, pero uno las tiene en parte porque las busca...

CW: Lo bonito de este libro es que prácticamente ustedes le cuentan al Dalai Lama, incluso muy didácticamente a veces, la ciencia, la historia de la ciencia en occidente, el tema epistemológico, en fin. Leyendo el libro, uno vuelve a descubrir la ciencia con asombro y lo que gusta mucho es que como que los científicos pueden distanciarse de la misma ciencia y mirarla desde fuera. Eso me parece importante.

FV: Lo es y ese es el desafío, porque tú no puedes tener un diálogo transcultural si sigues dogmáticamente anclado a tu disciplina como si fuera la pieza maestra de todo un discurso, eso obviamente no puede ser. Y esa primera vez tuvimos que invitar y escoger entre gentes que eran competentes en su materia, pero que tenían esa flexibilidad, o capacidad de ser abiertos. La otra cosa extraordinaria es que, habitualmente cuando se dan este tipo de diálogos —que ya se habían dado antes, cuando él viene a occidente—, son muy breves: se hace una mesa redonda, un par de personas dicen algo, él también, luego algunas preguntas diseminadas del público y se acabó. Aquí, en cambio, fueron ocho horas al día durante cinco días, en el living de su casa, sin periodistas, sin cámaras y, un hecho muy importante, dos traductores muy bien formados. Entonces, esa es la "receta" para que el diálogo pueda darse, porque si no, algo pasa, pero de manera muy superficial.

CW: Este tipo de diálogo uno lo hecha de menos en el mundo de hoy. Tal vez se deba a la especialización de la ciencia y de los saberes a causa de todos los desafíos que tiene el hombre de hoy. Los problemas que tiene le impiden darse el tiempo para abrir puentes ¿no es así? Se podrían, entonces, abrir tantos otros puentes así como los que se han tendido entre la ciencia y el budismo.

FV: Por cierto...

CW: El no hacerlo, ¿crees tú que es un problema que viene, además, de una manera de hacer ciencia o de entenderla?

FV: Sin duda, hay una contribución de la arrogancia implícita, un triunfalismo científico, pero a esa arrogancia de la ciencia ha contribuido también el público. Cuando ves que se le exige a la ciencia dar respuestas a todo y ser los pontífices de la verdad, se crea una

especie de tango a dos. Yo, en ese sentido, creo que es muy importante que la ética científica sea algo muy central, ser orgullosos de lo que podemos hacer, pero ser muy conscientes de sus límites, y, además, la ética del científico es la del ciudadano, no la del experto.

CW: Acabas de decir "la ética". En este libro planteas que habría dos miradas, una ética mas bién abstracta, teórica y otra, llamémosla encarnada, real, concreta. Si lo he entendido bien, significa eso, que hay una alternativa de ética a la ética que ha predominado en en occidente, una posibilidad de hacer otra ética y en tal caso ¿de dónde vendría esa posibilidad?

FV: Para mí esto es como el tema de las variaciones Goldberg. Algunas de las cosas que dijimos uno las puede declinar en muchos campos ¿no es cierto? Entonces en la ética encuentras reflejado el mismo problema que veíamos antes, es decir, que mientras yo siga pegado a la noción de que la ética es algo abstracto, de principios, de guías morales, lo que no es en sí mismo malo o negativo, pero tener una regla o entender lo que es un principio moral, no me cambia mi vivencia encarnada que tiene esa tendencia a alejarse. Cuando hablo ahí de la diferencia entre la ética abstracta y la ética moral y la ética encarnada y la pragmática, hago un llamado exactamente a ese mismo tipo de proceso de transformación que consiste en decir: puesto que hay una brecha entre el principio y mi comportamiento, no saco nada con seguirme repitiendo las mismas cosas. Cuando era niño me decían: "Hay que ser bueno, no hay que mentir", pero yo mentía igual y me sentía mal, porque no había que mentir. Nos ha pasado a todos ¿no? En cambio, en lo que toca a las prácticas de transformación, tengo que partir por aquellas experiencias que son espontáneas, que están situadas en el ser humano. De esas experiencias puedo partir y ampliarlas. Esas experiencias son el punto central del libro.

CW: ¿Puedes dar un ejemplo?

FV: En el libro doy un ejemplo tomado de la tradición taoísta, de Mentsu, un escritor chino. Es la parábola del pozo y dice: Si yo

veo a un niño que está parado al borde de un pozo y a punto de caer, no hay reflexión, no hay moral. No hay nada, no hay Kant, no hay Hegel: yo salvo al niño. Ese es el tipo de experiencia que presento en términos menos dramáticos, como cuando uno va a la oficina y de pronto hay un colega que manifiesta alguna dificultad, uno lo ve triste, preocupado, angustiado, deprimido, en fin, y en esos momentos, espontáneamente uno se siente tocado. Lo mismo ocurre cuando al hombre le es dado reflexionar; en lugar de dejar pasar ese momento como si fuera un microbús lo cultivamos, nos concentramos en eso, lo buscamos, encontramos maneras de enfocarlo, de ampliarlo, de estabilizarlo, y esa palabra *estabilidad* es clave. Es tomar esas capacidades espontáneas y cultivarlas, focalizarlas y partir desde ese punto en que yo no tengo que hacer ninguna reflexión.

CW: ¿Crees que es innato en el hombre sacrificarse por los demás o el hombre es más bien egocéntrico, un sobreviviente que necesita de todo un trabajo justamente para preocuparse del prójimo?

FV: Sacrificio es una palabra muy fuerte para esos ejemplos. Yo diría que es una simple capacidad de ser sensible, de tener una respuesta, porque no me estoy sacrificando por el otro al decirle "¿qué te pasa?"

CW: Pero si hay alguien que se está ahogando en un río, tú te tiras ¿Es un ejemplo válido?

FV: Sí pero no es el de Mentsú. En ese ejemplo, que es mucho más simple: veo al niño que se va a caer y no tengo una reflexión para ir en su ayuda, pero no me tiro necesariamente al pozo con él, eso es otro nivel de *get done and experiment.* El ejemplo de la oficina es parecido. No, el sacrificio es un tema particular. Yo diría que la capacidad de identificación, para ponerlo en los términos que me parece que corresponden aquí, con la del otro, el hecho de ponerme en el lugar del otro, que es una cosa muy común, es constitutivo al hombre y es demostrablemente falso por los propios criterios de la ciencia, que el hombre es intrínsecamente cerrado, egoísta y agresivo. Hoy por hoy —y te voy a dar un solo elemento de evidencia empírica— hay una cantidad importante de estudios sobre primates superiores. Durante

muchos años se pensaba que los primates superiores —los gorilas, los chimpanzés, etc.— eran seres muy agresivos, muy individualistas. Después de cuarenta años de estudios etológicos sobre los primates superiores, la conclusión está así como decantada. ¿Qué es un primate superior? ¿Es aquel que tiene dos brazos, que tiene nariz, que tiene dos ojos? No, el primate superior tiene una cosa fundamental y que lo hace primate superior y que lo identifica con el hombre, y es que es un experto en el estado de ánimo del otro.

Si hay algo que todos los primates tienen en común es el poder colocarse en el lugar del otro. En el fondo, el primate, como el hombre, es un ser que está parado en el mundo de tal manera que el estado del otro lo interpela a fondo.

CW: ¿Eso no lo tiene ningún otro animal? Pienso en los delfines.

FV: Los delfines, tal vez; los cetáceos son asunto aparte, pero ciertamente entre los primates inferiores, no es así. Somos intérpretes constantes del estado mental, anímico del otro y basta pensar en los niños, en la vida materno-infantil. El niño y su madre es una cosa increíble de tú-y-yo y yo-y-tú, sin mayor barrera, y esa es una lección que tiene que ser muy bien escuchada. Hay muchos estudios recientes que demuestran eso. Entonces la capacidad empática, la capacidad de ponerse en la posición del otro es una cuestión que tenemos en el fondo de nuestros genes.

CW: ¡Esa es una muy buena noticia!

FV: — Efectivamente, pero eso no quiere decir que las tendencias a defenderse, a cerrar la territorialidad no existan ni que esas tendencias sean espontáneamente desarrolladas. Requieren justamente ser cultivadas y eso es lo que los primates no hacen.

CW: ¿Cuáles son las implicancias concretas de esta nueva ética?

FV: Si mi declaración es que me interesa transformarme en un ser humano más allá del que soy en la cotidianidad y si tengo un compromiso con el otro y un compromiso a hacer una contribución, entonces tengo que comprometerme con un camino de transformación.

CW: ¿Es una ética más realista de lo que es el ser humano?

FV: Mucho más realista, es la única ética realista en la medida en que es la única que te da un fundamento, un quehacer que te lleva a un cambio.

CW: ¿La otra ética estaría condenada siempre al fracaso?

FV: Una ética condenada a repetir los mismos errores del pasado. Se ha incendiado, quemado y torturado en nombre de los principios de la moral hasta el día de hoy. Entonces sí, tengo cierta distancia...

CW: Una de las grandes preocupaciones de occidente es el yo. Siempre estamos diciendo yo, en todo. "Yo soy el que soy", decía el Quijote en el momento en que lo tiran del caballo y pataleaba

FV: (*sonríe*) Si me hubieras hecho esta pregunta la semana pasada, te habría invitado a la conferencia que di en el congreso de psiquiatría y neurología sobre la naturaleza del yo cognitivo. Por el momento saquemos el budismo.

CW: Muy bien.

FV: Dejémoslo tranquilo un momentito. Porque es una de las cosas fascinantes, que en algún momento la gente tendrá que hacerse cargo de ellas, ¡son noticias del reino de la ciencia!

CW: Del pensamiento humano.

FV: Como tú dices, de la filosofía en acción que trae la ciencia. En las ciencias cognitivas más duras, la neurociencia, la cirugía experimental, la inteligencia artificial, las cosas más aterrizadas de la ciencia, se descubre un yo cognitivo, es decir, una persona que es capaz de ver, pensar, reconocer, moverse, tener planes, e tener memoria etc. Pero cuando uno lo somete al análisis, uno se encuentra —y eso es lo fascinante—, no con el yo, sino con los pedazos que hacen memoria o que son capaces de movimiento, o que son sensitivos a los bordes en movimiento. Hay, lo que en ciencias cognitivas o en neurofisiología se llaman los módulos cognitivos, y esos módulos son incluso identificables a nivel de

subcircuitos neuronales y tienen transmisores propios y yo puedo hacer modelos en robots, o sea, una cuestión superconcreta. Entonces, si tú me preguntas dónde está el yo, te respondo que no está aquí ni aquí ni acá.

CW: Está diseminado...

FV: Mejor que eso, y ahí es donde la ciencia aporta un elemento de respuesta que es, a mi entender, una de las cosas más creativas del siglo veinte, que es entender lo que es un fenómeno. A mí me gusta designar por el nombre de "emergencia"o de autoemergencia o de autoorganización, es decir, que estos pedazos interactúando entre sí, bajo ciertas reglas, bajo cierta legalidad, dan origen a un conjunto armónico bajo procesos que uno puede explicitar muy concretamente, o sea, la vieja idea de que el todo es la suma de las partes, pero aquí la suma de las partes, sabemos que no es la suma en realidad, sino que es, para decirlo técnicamente, un interacción no lineal de las partes y ahí es donde viene todo el tema de las matemáticas, de la no linealidad y la complejidad, y el caos juega ahí un rol central. Es decir que yo puedo escribir ecuaciones para mostrar cómo se pasa de lo uno a lo otro sin contradiccción. Entonces ese yo, que aparece así, bajo ese mecanismo de autoemergencias no lineales, tiene —y esto es el *punch line*, como dicen los norteamericanos— una manera de ser muy curiosa porque tiene una identidad en el sentido de que tiene cierta coherencia que me permite tomar vasos, decir hola, venir al programa, hablar en la televisión y al mismo tiempo necesito ser presentado como Francisco Varela, que a la vez no existe, porque yo lo trato de encontrar y no está, entonces es un modo de existencia muy peculiar, que está justo entre la existencia y la inexistencia...

CW:¿Cómo es eso?

FV: Pero esto no es privativo de nosotros, en la naturaleza también ocurre, no somos sustanciales ni somos inexistentes y el budismo hace exactamente el mismo descubrimiento con su fenomenología. ¡Sostener que el budismo dice que no existimos, es una aberración! El budismo pone su dedo exactamente en la misma cosa. Por ejemplo, el Dalai Lama usa una expresión coloquial tibetana muy bonita al

hablar de lo que es el yo. Dice: *gzugzu gzigzu windu* (*apoya el índice sobre la mesa*), donde *gzugzu*, quiere decir "dedo", *gzigzu* y la partícula negativa *windu*, significa: "no encuentro dónde", entonces el yo es aquello donde yo no encuentro dónde poner el dedo.

CW: No puedo dejar de hacerte la última pregunta. Después de este libro vinieron varios encuentros más. En uno de ellos se trató el tema del sueño y de la muerte. Aquí no podemos hablar de la muerte tan científicamente, pero sí con intuiciones. Se ve que eres un hombre intuitivo, que vas más allá de lo que simplemente vas demostrando; por lo menos, es la sensación que yo tengo.

FV: Si no es así, no puedes hacer ciencia.

CW: Creo lo mismo. ¿Qué intuición tienes respecto del yo, la conciencia y la muerte? ¿Hay una respuesta válida ahí?

FV: Todas las preguntas son válidas. Una primera y rápida reflexión la encontramos en un libro que acaba de aparecer en inglés, que yo también edité, y que se llama: *Dormir, soñar y morir*, (*Sleeping, Dreaming and Dying**), que está exactamente en el mismo formato y sobre esos temas. Y efectivamente ahí se tocan estos temas Mi concepto para hacer esa reunión, que es la cuarta que tuvimos con el Dalai Lama, fue tocar aquellas zonas que yo llamo las "zonas en penumbra del yo", en que uno se duerme, en que uno se va, en que uno no está, en que uno se muere... ¿Y qué pasa allí, cuál es mi intuición? Mi intuición es que efectivamente es una arrogancia terrible pensar que la conciencia y lo mental son una propiedad individual. No es verdad, ya que, desde el punto de vista de lo que hablábamos, de mi mente y tu mente, mi mente y el mundo, son cosas que están íntimamente implicadas. Además, cuando uno mira la vida más a largo plazo, en una perspectiva histórica y temporal mucho más larga, uno se da cuenta de que lo que los budistas describen como la continuidad de la conciencia no es tan distinto de lo que uno dice cuando mira a su alrededor. Por

1. *Dormir, soñar, morir. Nuevas conversaciones con el Dalai Lama.* F. Varela, editor. Dolmen Ediciones, Santiago de Chile, 1998. Libro traducido y publicado posteriormente a esta entrevista.

ejemplo, hoy en la mañana miraba una higuera muy bonita que tienen mis padres en su casa, aquí en Chile, y veía los higos que comenzaban a salir. Me dije que tranquilamente podría pensar que esos higos son la reencarnación de los higos del año pasado... Es decir que hay leyes de recurrencia a largo plazo, que hace que haya cosas que permanecen pero que no son individuales. El higo no es el mismo, pero esa "higuedad" se re-constituye. Creo entonces, y esa es mi intuición, que en la muerte hay una desaparición de este yo individual, pero que eso no es el final, sino que hay unas corrientes de causalidad. Y hablo de causalidad porque no son cosas para mí misteriosas, es parte de la legalidad del mundo, que va más allá, que tiene una continuidad de otra temporalidad, de otros ciclos temporales, y no le atribuyo a eso un valor trascendente, lo considero más bien un misterio de la profundidad de lo natural.



Francisco Varela se doctoró en Biología en la Universidad de Harvard, en 1970. Desde entonces ha enseñado y dirigido investigaciones en diversas universidades, incluyendo la Universidad de Colorado en Boulder, la Universidad de Nueva York y el Instituto Max Planck (Alemania) para la investigación Cerebral. En la actualidad es director de Investigación en el Centro Nacional de Investigación Científica en París.

Varela es autor de varios libros, entre los que destacamos *El árbol del conocimiento*, en coautoría con Humberto Maturana (Ed. Universitaria), *The Embodied Mind: Cognitive Science and Human Experience (La mente encarnada: ciencia cognitiva y experiencia humana*, en colaboración con E. Thompson y E. Rosch, que ha sido traducido a ocho idiomas , MIT Press, 1991), y *Etica y Acción* (Dolmen ediciones, 1995). También son de su autoría más de 200 artículos para publicaciones científicas en las áreas de la neurociencia y la ciencia cognitiva.

En 1987 fundó las Conferencias « Mente y Vida », una serie de diálogos entre científicos

y el Dalai Lama. Dolmen Ediciones ha publicado la version española de dos libros resultantes de dichos encuentros : Un Puente para dos Miradas : Conversaciones con el Dalai Lama Sobre las Ciencias de la Mente (1997), seguido de Dormir, Soñar, Morir : Nuevas conversaciones con el Dalai Lama (1999).

Otros títulos publicados por el autor en Dolmen Ediciones:

- Ética y acción (1997).
- Un puente para dos miradas. Conversaciones con el Dalai Lama sobre ciencias de la mente, con *Jeremy Hayward* (1998).
- Dormir, soñar, morir. Nuevas conversaciones con el Dalai Lama (1999).



"Lo que he querido desarrollar en estos ensayos es una comprensión del fenómeno de la vida en toda su gloria y majestad. Esto quiere decir evitar caer en la tentación dominante de partir del universo impersonal y muerto de la física para luego verse forzado, casi a contrapelo, a rescatar de ahí lo vivo, cuya especificidad es lo individual creador de significado. Gran parte de las ideas que aquí presento son precisamente el esfuerzo por romper ese sueño newtoniano, y por recentrar el fenómeno de la vida sobre sus propias bases, es decir, no como evento improbable que escapa a la muerte física, sino como ocurrencia central de la naturaleza que funda todo conocer y, en particular, el conocimiento del mundo físico. La conclusión a la que estos ensayos apuntan es que lo central es un circularidad inalienable entre el acto de conocer y vivir, entre el universo de lo vivo y el conocer como objeto de estudio. Dicho de otra manera, el fenómeno de la vida, como un todo, quiere decir, precisamente, que el acto de vivir precede a la explicación del origen de la vida sobre la Tierra. Que el conocer precede a la comprensión del conocer visto como mecanismo biológico y neuronal. Que la experiencia vivida es la base misma de la exploración científica de la concienc

F. V.





